Stochastic parametrization of ocean models through high resolution observations

Francesco Tucciarone

COLLEGE MATHS, TELECOMS
DOCTORAL INFORMATIQUE, SIGNAL
BRETAGNE SYSTEMES, ELECTRONIQUE

Université de Rennes

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES

ÉCOLE DOCTORALE N° 601
Mathématiques, Télécommunications, Informatique, Signal, Systèmes, Électronique
Spécialité : *Informatique*

Par
Francesco Leonardo TUCCIARONE

Stochastic parametrization of ocean models through high resolution observations

Rapporteurs avant soutenance :

Julie DESHAYES Directrice de recherche, CNRS, IPSL Paris
Laurent DEBREU Directeur de recherche, INRIA, Centre U. Grenoble

Composition du Jury :

Président :	Nicolas CROUSEILLES	Directeur de recherche, INRIA, Centre U. Rennes
Examinateurs :	Quentin JAMET	Researcher, SHOM, Brest
	Patrick MARCHESIELLO	Directeur de recherche, IRD, Toulouse
Dir. de thèse :	Etienne MÉMIN	Directeur de recherche, INRIA, Centre U. Rennes
Co-dir. de thèse :	Long LI	Starting Research Position, INRIA, Centre U. Rennes

List of contributions

Conference presentations

— Stochastic Modelling and Location Uncertainty (LU) formalism: hydrostatic Boussinesq equations and their implementation in NEMO, 23rd May 2022, EGU 2022 General assembly, Vienna, Austria, Online presentation.

— Data driven stochastic primitive equations with dynamic modes decomposition, 24rd April 2023, EGU 2023 General assembly, Vienna, Austria, Online poster.

Conference papers

— Francesco L. Tucciarone, Etienne Mémin, Long Li (2022). Primitive Equations Under Location Uncertainty: Analytical Description and Model Development. In: Chapron, B., Crisan, D., Holm, D., Mémin, E., Radomska, A. (eds) Stochastic Transport in Upper Ocean Dynamics. STUOD 2021. Mathematics of Planet Earth, vol 10. Springer, Cham.

— Francesco L. Tucciarone, Etienne Mémin, Long Li (2023). Data driven stochastic primitive equations with dynamic modes decomposition. In: Chapron, B., Crisan, D., Holm, D., Mémin, E., Radomska, A. (eds) Stochastic Transport in Upper Ocean Dynamics. STUOD 2022. Mathematics of Planet Earth, vol 10. Springer, Cham.

— Francesco L. Tucciarone, Etienne Mémin, Long Li (Submitted). Data driven stochastic primitive equations with dynamic modes decomposition. In: Chapron, B., Crisan, D., Holm, D., Mémin, E., Radomska, A. (eds) Stochastic Transport in Upper Ocean Dynamics. STUOD 2023

Journal papers

— Francesco L. Tucciarone, Long Li, Etienne Mémin, Pranav Chandramouli (in preparation). Stochastic data-driven model of sub-mesoscale eddies in gyre circulation. In: Journal of Advances in Modeling Earth Systems (JAMES)

Adieu, Pannychis;
mais ne crois pas que nous échapperons l'un à l'autre. J'ai voulu asservir le monde à ma raison et, dans cet antre humide, je me suis heurté a toi. Toi, le monde, tu essayais de le soumettre à l'empire de ton immagination. Et ainsi en ira-t-il pour tous les temps à venir; deux camps adverses se feront face. Les uns verront dans le monde un système, un ordre, alors qu'aux autres, il apparaîtra comme un monstre. Les premiers diront qu'il faut remettre le monde en cause et affirmeront qu'il est possible de le modeler comme une pierre en la taillant. Quant aux seconds, convaincus que le monde se modifie de lui-même dans toute son opacité, tel un monstre changeant sans cesse d'allure, ils l'accepteront tel quel : pour eux, on ne peut remettre le monde en cause que dans l'infime mesure où la couche extrêmement fine de l'entendement a un pouvoir sur les prodigieuses forces tectoniques des instincts humains. Alors que le camp des uns se verra reprocher son pessimisme par les autres, ceux-là passeront pour se nourrir d'utopies aux yeux de ceux-ci. L'Histoire obéit à certaines lois, affirmera-t-on ici, à quoi d'autres répondront que ces lois ne sont que les produits de l'esprit des hommes. Notre conflit, Pannychis, celui qui oppose le devin à la Pythie, s'enflammera dans tous les domaines, mais il est encore chargé d'émotions et peu réfléchi; pourtant, voilà déjà que l'on construit un théâtre et qu'à Athènes, un écrivain inconnu compose une tragédie d'Œdipe. Cependant, Athènes est une ville de province et Sophocle tombera dans l'oubli, alors que l'histoire d'Œdipe subsistera et nous posera des énigmes.
Maintenant, à quelle force Œdipe doit-il son sort ? Aux dieux? Ou bien lui a-t-il été infligé pour avoir transgressé certaines lois qui constituaient le socle de la société d'alors — ce que j'ai voulu l'empêcher de faire en dictant mon oracle? Ou Œdipe aurait-il encore été victime d'un hasard conjuré par les caprices de tes vaticinations?

Friederich Dürrenmatt
La mort de la Pythie, 1976

Table of Contents

General introduction

Context and motivation

The global climate strongly depends on the Ocean's state. The interaction with the Atmosphere, in the form of mutual energy fluxes exchanges and heat redistribution, plays a crucial role in the climate regulation. The understanding of the current state of the global Ocean has greatly increased with the aid of observations and measurements, but numerical simulations remain the only way to forecast the system and assess future states. This is fundamental for predicting meteorological or climatological related hazards. The primary tool for the investigation of the Ocean and the Atmosphere are large-scale simulations, while high resolution simulations remains confined to small geographical domains or short integration periods. The complex interdependence of mesoscale and sub-mesoscale dynamics is, however, lost in simulations that do not resolve scales below the Rossby radius of deformation, and thus must be parametrised. Most of the challenges of fluid dynamics (in all its connotations) arise from the representation of these effects with an efficient closure scheme. A novel research trend involves incorporating perturbations and noise components into the dynamics. The goal is to enrich the variability and parametrise sub-grid processes, turbulence, boundary value uncertainty and account for numerical and discretization errors.

Aim and objectives

This work addresses the benefits of a stochastic framework where the Lagrangian trajectory is decomposed into a smooth-in-time large scale velocity and a random, fast-evolving uncorrelated part. This approach, named Location Uncertainty (LU), is built upon a stochastic version of the Reynolds' transport theorem that allows to cast the classical physical conservation and balance statements into a framework handling a scale separation between fast and slow scales. Ideally, the separation of scales enhances the representation of the subgrid processes. This framework has already been tested and was proven successful in several configurations with different underlying geophysical models. The hy-

drostatic primitive equations are the main target of this manuscript. Their importance in the representation fo the ocean is discussed and the application of the LU framework to these equations is assessed within the community model NEMO (https://www.nemo-ocean.eu).

Structure of the manuscript

Chapter 1 provides a primer in fluid mechanics, describing the fundamental equations of mechanics and thermodynamics of a flow. These general equations are then specialised to the case of an environmental flow and simplified enough to allow the building of an Oceanic General Circulation Model, aiming at simulating the large scale circulation. The main equations, variables of interest and critical points of an OGCM are presented to provide an overview of the complexity of such tools. Chapter 2 introduces the main framework of this thesis, the Location Uncertainty principle. This framework has been proposed by Etienne Mémin in 2014 and has already been tested in several reduced model, such as Surface Quasi-Geostrophic equations (SQG), Quasi-Geostrophic equations (QG) and Rotating Shallow-Waters (RSW). The development followed in the exposition aims at introducing all the models that must be considered in the development of a stochastic ocean general circulation model. Chapter 3 describes the different models that have been adopted to generate the noise for the stochastic parametrization proposed in the Location Uncertainty framework. Finally, chapter 4 summarizes the numerical results obtained with the location uncertainty framework implemented in the state of the art OGCM NEMO.

Introduction (Français)

Le climat, la météo et l'océan (avec ses mers marginales) sont les trois acteurs qui ont joué le plus grand rôle dans l'évolution de l'humanité telle que nous la connaissons. Les établissements humains, migrations et les découvertes ont souvent été motivées par un facteur climatique ou météorologique. Les événements cataclysmiques et leurs conséquences sur les mers ont façonné les récits épiques et les croyances religieuses. Les interactions avec la mer et l'océan ont caractérisé la société, à tel point que les rituels apotropaïques impliquant la mer sont encore présents dans de nombreuses cultures. En outre, l'importance de l'océan a été légalement reconnue par la Convention des Nations Unies sur le droit de la mer de 1982, qui établit que la gestion rationnelle des ressources marines et leur conservation pour les générations futures sont primordiales. Dans l'ombre du changement climatique, les sciences de l'atmosphère, du climat et de l'océan sont fondamentales pour donner un aperçu de ce que l'avenir peut nous réserver, les prévisions météorologiques, océaniques et climatiques numériques constituant une base solide pour l'évaluation des risques à court et moyen terme et la discussion sur les politiques d'atténuation.

Les prévisions météorologiques ont longtemps été basées sur les connaissances et l'expérience des prévisionnistes plutôt que sur une application rigoureuse de la théorie des systèmes dynamiques (voir par exemple Smagorjnsky, 1983). Selon Kalnay (2012), la première intuition dans le développement de la prévision météorologique numérique fut celle du physicien norvégien V. Bjerknes en 1904, qui reconnut que l'état futur de l'atmosphère est, en principe, complètement déterminé par son état initial précis et les conditions aux limites connues, ainsi que par les équations du mouvement de Newton, l'équation d'état de Boyle-Charles-Dalton, l'équation de continuité de la masse et l'équation de l'énergie thermodynamique. La première tentative d'évaluation de l'évolution météorologique fut réalisée par Lewis Fry Richardson en 1922, en intégrant manuellement les équations d'évolution avec un pas de temps de 6 heures. L'échec de ce projet est lié à l'état de l'art qui prévalait à l'époque : c'est-à-dire, d'une part, il était impossible ou improbable de recruter les 64 000 personnes estimées nécessaires pour atteindre les résultats souhaités. D'autre part, la condition de Courant-Lewy-Friedrich pour la stabilité des systèmes hyperboliques n'avait pas encore fait son apparition. En fait, ce n'est qu'en 1928

que Courant-Lewy-Friedrich proposeront cette théorie (Lewy et al., 1928). Alors que les tentatives de prévisions météorologiques numériques ont repris à la fin des années 40 et au début des années 50 avec les travaux fondamentaux de von Neumann et de G. Charney (voir, par exemple, Charney, 1951). Ces intuitions réussies sur la possibilité de prévoir le temps ont ensuite été appliquées à d'autres fluides géophysiques, tels que l'océan. Dans ce contexte, les travaux de Bryan (Bryan, 1967) apparaissent comme la première tentative de simulation de l'état de l'océan. Le premier modèle de circulation générale de l'océan avec des géométries réalistes fut proposé par Bryan and Cox en 1968 (Bryan and Cox, 1968). Ce travail a ouvert la voie à ce qui est aujourd'hui l'un des principaux paradigmes en océanographie, le modèle Bryan-Cox-Semtner (Semtner et al., 1988; Dukowicz et al., 1994; S. M. Griffies, Böning, et al., 2000; Murray et al., 2001). Les premières tentatives de couplage des modèles atmosphériques et océaniques furent entreprises par Manabe et al. (1969) établissant le paradigme des modèles de systèmes terrestres couplés. Malgré cet intérêt soutenu et toujours croissant pour le développement d'outils de modélisation des océans et du climat, ainsi que l'augmentation des capacités de calcul, un problème fondamental demeure : les équations fondamentales de la mécanique des fluides présentent un vaste ensemble d'interactions qui sont par nature impossibles à capturer complètement avec des simulations. Les équations de mouvement d'un fluide (voir le Chapitre 1 de cette thèse) suivent une structure non-linéaire qui peut être étudiée dans un premier temps à travers une simple équation d'advection:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{1}$$

Si l'on attribue à cette équation une condition initiale analytique telle que $u(x, t_0) = A\cos(kx)$, on peut déduire, par expansion de Taylor, l'état du système proche de l'état initial $t = t_0$. Ou bien,

$$\begin{aligned} u(x,t) &= u(x,t_0) + (t-t_0)\frac{\partial u(x,t_0)}{\partial t} + \mathcal{O}(\Delta t) \\ &= A\cos(kx) - (t-t_0)\,u\frac{\partial u(x,t_0)}{\partial x} + \mathcal{O}(\Delta t) \\ &= A\cos(kx) - (t-t_0)\frac{A^2k}{2}\sin(2kx) + \mathcal{O}(\Delta t)\,. \end{aligned}$$

Dans la dernière ligne, on remarque que la solution a une fréquence harmonique de $2k$, deux fois plus élevée que celle de la condition initiale. Un mécanisme de transfert de la

grande échelle à la petite échelle est donc encodé dans le terme non linéaire, qui créera des échelles de plus en plus petites à mesure que le mouvement persiste. Des résultats similaires s'appliquent aux mouvements bidimensionnels ou tridimensionnels (Bailly et al., 2015). Cette cascade d'énergie et d'échelle se termine à une échelle spécifique, l'échelle de dissipation de Kolmogorov (Kolmogorov, 1941), où la friction moléculaire arrête le mouvement, convertissant l'énergie cinétique en énergie thermique par le biais de la chaleur de Joule. Toute résolution numérique visant à être pleinement représentative de l'écoulement doit être effectuée jusqu'à l'échelle de Kolmogorov, de l'ordre de $10^{-6}m$. La réalisation de cette tâche pour un écoulement fluide avec un nombre de Reynolds Re donné nécessite (dans un cadre très optimisé) la capacité de gérer $N^3 > \mathrm{Re}^{9/4}$ points de collocation spatiaux (Bailly et al., 2015). En plus de cette limitation inhérente de la résolution, une autre limitation "structurelle" est présente dans les modèles de circulation générale océanique (en anglais, ocean general circulation model ou OGCM). La résolution numérique d'une équation aux dérivées partielles commence toujours par la discrétisation des fonctions inconnues, la discrétisation de l'opérateur différentiel ou une combinaison des deux. Les méthodes spatiales de Galerkine sont un exemple de la première approche, où la fonction inconnue est représentée comme une combinaison de fonctions continues par parties d'un certain ordre (Zienkiewicz et al., 2013). Les méthodes des différences finies sont un exemple de la seconde approche, où l'opérateur différentiel continu est remplacé par un opérateur discret opérant sur un ensemble de points discrets. Les deux approches sont souvent combinées, par exemple la progression temporelle d'une méthode Galerkine spatiale avec un schéma de différence finie dans le temps. Les approximations par différences finies des opérateurs différentiels continus sont généralement basées sur l'expansion de Taylor et leur degré de précision dépend de la troncature adoptée. Il s'ensuit qu'une augmentation de la résolution du modèle discret augmente également la fidélité de l'opérateur discret. Une tentative naïve de s'attaquer au problème consisterait à d'espérer l'augmentation des ressources informatiques, afin de permettre des calculs de plus en plus importants. Cette tâche n'est toutefois pas envisageable dans un avenir proche : la plus petite échelle de temps dans l'océan peut être considérée comme $1s$, tandis que les plus petites échelles spatiales peuvent être considérées comme $10^{-3}m$. Une simulation climatique (d'une durée $\sim 1000y$) de l'océan nécessiterait $\sim 10^{27}$ cellules de grille (S. M. Griffies and Treguier, 2013). De plus, même si la troncature est adoptée à une certaine échelle, comme celle donnée par le rayon de déformation de Rossby (aussi appelée méso-échelle), la prise de conscience de l'importance de la dynamique sous-méso-échelle pour les flux océaniques

est bien documentée par une série de publications récentes (Capet et al., 2008; Callies et al., 2015; Lévy, Klein, et al., 2010; McWilliams, 2016). Ces tourbillon des sous-méso-échelles, typiquement de l'ordre de $\mathcal{O}(1)$ à $\mathcal{O}(10)$ kms, ont été observgiés comme jouant un rôle vital dans le soutien des courants énergétiques à méso-échelle qui sont d'un ordre de grandeur plus grand à l'échelle – $\mathcal{O}(100)$ km. Ils jouent également un rôle essentiel dans la caractérisation de la circulation océanique à grande échelle, le mélange des traceurs dans les couches supérieures de l'océan et la cascade énergétique (Capet et al., 2008). Leur présence est observée par des mesures de terrain satellitaires (Durand et al., 2010) et leur importance est soulignée par des études numériques (Callies et al., 2015; Della Penna et al., 2019). Leur contribution aux échanges biogéochimiques de matière dans l'océan est mise en avant dans les travaux de Capet et al. (2008) et Lévy, Iovino, et al. (2012). La résolution explicite de ces courants à sous-méso-échelle dans les simulations numériques est un défi informatique qui nécessite des résolutions plus petites que le plus petit tourbillons à résoudre (B. Fox-Kemper et al., 2019). Cela nécessite des dimensions de grille inférieures à 1 km, par rapport aux dimensions de grille de 10 à 100 km généralement utilisées dans les modèles de circulation générale des océans. Il est intéressant de noter que la résolution explicite des tourbillons à sous-méso-échelle est possible dans des scénarios limités dans lesquels il a été démontré qu'ils avaient un impact significatif sur les caractéristiques de l'écoulement (Lévy, Klein, et al., 2010; Hurlburt et al., 2000; Siegel et al., 2001). La modélisation à haute résolution nécessite cependant des considérations plus fines pour les méthodes numériques employées, posant de nouveaux défis aux modèles océaniques. L'extension du spectre d'application de ce type de modèle, initialement développé pour des configurations basse résolution à grande échelle, pose de nouveaux défis de nature numérique et physique (S. M. Griffies and Treguier, 2013) et nécessite la représentation précise d'un spectre d'énergie plus large (Lemarié et al., 2015). Une façon de résoudre ce problème est de recourir à des techniques d'imbrication, où une grille à résolution relativement grossière couvre l'ensemble du domaine de calcul et est affinée localement dans les zones d'intérêt telles que les côtes, les détroits ou les zones d'activité intense connues. Cette approche multi-résolution est suffisamment avancée/solide (Debreu and Blayo, 2008; Debreu, Marchesiello, et al., 2012) pour être utilisée afin d'atteindre localement des résolutions raffiné qui résolvent les sous-méso-échelles (Marchesiello et al., 2011). Des stratégies multi-résolution et multi-grille sont employées pour comprendre la dynamique des traceurs passifs (Jouini et al., 2013; Bricaud et al., 2020). Plusieurs modèles de circulation générale mettent aujourd'hui en œuvre la stratégie bien connue de

raffinement adaptatif des grilles de l'AGRIF (Debreu and Blayo, 2002) (voir par exemple Madec et al., 2019; Auclair et al., 2022). Lorsque la résolution augmente, un traitement plus adéquat du transfert d'énergie entre les échelles résolues et non résolues est nécessaire (Sanderson, 1998; Thuburn et al., 2014), tandis que le mélange dianeutre fallacieux reste un aspect complexe en présence de tourbillons à méso-échelle (Ilıcak et al., 2012). Malgré les nouveaux défis posés par une numérisation plus sophistiquée pour un modèle à haute résolution, les modèles côtiers sont des outils de recherche très répandus. Ces codes diffèrent si fortement des modèles à grande échelle que leur mise en œuvre est souvent une variante autonome d'un modèle de circulation générale (voir par exemple les différences entre CROCO et ROMS) soit un projet complètement indépendant (voir par exemple Delft3D, MARS Lazure et al., 2008) avec des approches éventuellement complètement différentes (Defina, 2007). Le développement de ces outils est principalement motivé par la nécessité de préserver les communautés côtières. Il n'est donc pas rare de voir ces modèles employés dans des études d'influence de la marée pour évaluer l'éventuel risque combiné de marée et de surcote (Dumas et al., 2012; Idier et al., 2012) là où les marées jouent un rôle important (par exemple dans la Manche) ou là où les ondes de tempête sont particulièrement fréquentes (par exemple dans le nord de l'Adriatique [1] , cf. Mel et al., 2013; De Zolt et al., 2006). L'introduction d'un forçage stochastique constitue une alternative intéressante pour modéliser les contributions des tourbillons. Par exemple, Porta Mana et al. (2014) ont appliqué une paramétrisation stochastique des tourbillons à méso-échelle basé sur des variables résolues pour une configuration quasi-géostrophique de double-gyre avec des résultats encourageants. Franzke et al. (2005) proposent un cadre stochastique différent pour un modèle barotrope réaliste en définissant des modèles d'ordre réduit. Un modèle de fermeture stochastique non linéaire et semi-analytique a été proposé par Grooms and A. J. Majda (2013) dans le cadre de la super paramétrisation de Grabowski et al. (1999). Cependant, les travaux de Berloff (2005) suggèrent qu'un forçage aléatoire ne rend pas toujours compte de toutes les interactions non linéaires fondamentalement importantes aux échelles résolues. En outre, ces forçages aléatoires ne sont pas contraints par des invariants physiques et ne sont par conséquent pas toujours physiquement pertinents. Dans cette optique, deux méthodologies complémentaires ont été introduites par Mémin (2014) et Holm (2015), fournissant une méthodologie rigoureusement justifiée pour définir des représentations stochastiques à grande échelle des équations de Navier-Stokes

1. Le lecteur à l'aise avec un livre en italien pourrait être intéressé par "Un giorno, ospite inatteso, arrivò l'alluvione" du professeur émérite Luigi D'Alpaos, qui décrit deux inondations extrêmes dans le nord de l'Adriatique et leur impact sur les communautés.

(Debussche et al., 2023) conservant l'énergie et la circulation, respectivement. Ces deux modèles reposent sur une décomposition stochastique de la trajectoire lagrangienne en une composante lisse dans le temps induite par la vitesse à grande échelle et un bruit aléatoire de déplacement en évolution rapide non corrélé. Le contexte théorique solide permet de définir une représentation à grande échelle avec une composante stochastique représentant la contribution sous-grille, introduisant des degrés de liberté supplémentaires à exploiter dans la modélisation de phénomènes spécifiques, tels que les composantes à grande échelle (Tucciarone et al., 2022; Tucciarone et al., 2023), la turbulence à petite échelle (Harouna et al., 2017; Chandramouli, Mémin, and Heitz, 2020), les effets de couche limite (Pinier et al., 2019), les processus de convection (Jamet et al., 2023) ou pour concevoir des modèles intermédiaires (Harouna et al., 2017; Chapron et al., 2018; Pinier et al., 2019; Bauer et al., 2020a; Cintolesi et al., 2020a). Le modèle Location Uncertainty (LU) (Mémin, 2014) a été appliqué au modèle quasi-géostrophique barotrope (Bauer et al., 2020a; Bauer et al., 2020b), au modèle quasi-géostrophique barocline (Li, Deremble, et al., 2023) au modèle d'eau peu profonde (Brecht et al., 2021), au modèle quasi-géostrophique de surface (Resseguier, Li, et al., 2021), aux équations primitives hydrostatiques (Tucciarone et al., 2022; Tucciarone et al., 2023) et a récemment été formulé pour les équations de Boussinesq non hydrostatiques (Tissot, Mémin, et al., 2023), prouvant son efficacité pour structurer l'écoulement à grande échelle (Bauer et al., 2020a), reproduire les statistiques à long terme (Bauer et al., 2020b) et pour fournir un bon compromis entre la représentation de l'erreur du modèle et la propagation de l'ensemble (Resseguier, Li, et al., 2021; Brecht et al., 2021). Dans ce travail, le cadre stochastique physiquement cohérent de Location Uncertainty (Mémin, 2014; Resseguier, Mémin, and Chapron, 2017a) est appliqué à une configuration idéalisée/optimisée de double-gyre en 3D en utilisant un modèle piloté par les données pour le bruit afin de caractériser les contributions tourbillonnaires non résolues. Le double-gyre est une configuration idéale pour cette étude car elle montre des variations significatives dans les statistiques d'écoulement, même dans les champs moyens, lorsque la résolution horizontale passe d'une résolution tourbillonnaire à une résolution sous-méso-échelle (Lévy, Klein, et al., 2010).

INTRODUCTION (ENGLISH)

The Climate, the Weather and the Ocean (with its marginal seas) are the three players that have taken the biggest role in shaping human evolution as we know it. Human settlements, migrations and discoveries have often been driven by a climatic or meteorological forcing. Cataclysmic events and their consequences on the seas shaped epic narrations and religious faiths. The interaction with the sea and the ocean has been characterising the society so much so that apotropaic rituals involving the sea are still present in many cultures. Moreover, the importance of the ocean was legally recognised with the 1982 United Nations Convention on the Law of the Sea, establishing that the rational management of marine resources and their conservation for future generations is paramount. In the shadow cast by climate change, atmospheric, climate and ocean sciences are fundamental to offer insights into what the future may awaits us, with numerical weather, ocean and climate prediction as solid basis for short and medium term risk assessment and for discussions on mitigation policies discussion.

Weather prediction has been for a long time based on the knowledge and experience of the forecaster rather than a rigorous application of dynamical system theory (see *e.g.* Smagorjnsky, 1983). According to Kalnay (2012), the first intuition in the development of numerical weather prediction was by Norwegian physicist V. Bjerknes in 1904, who recognized that the future state of the atmosphere is, in principle, completely determined by its detailed initial state and known boundary conditions, together with Newton's equations of motion, the Boyle–Charles–Dalton equation of state, the equation of mass continuity, and the thermodynamic energy equation. The first attempt at evaluating the weather evolution was performed by Lewis Fry Richardson in 1922, integrating by hand the evolution equations with a time step of 6-hours. The failure of this project was related to the state of the art prevalent in those years: that is, on the one hand, the estimated 64000 workforce people anticipated to achieve the desired results were impossible/unlikely to be recruited. On the other hand, Courant-Levy-Friedrich's condition for stability of hyperbolic systems had yet to make an appearance. In fact, it was not before 1928 that Courant-Lewy-Friedrich would propose the theory (Lewy et al., 1928). Whereas attempts in numerical weather predictions would restart in the late 40s and early 50s with the sem-

inal work of von Neumann and G. Charney (see *e.g.* Charney, 1951). Subsequently, these successful intuitions about the possibility of forecasting the Weather were applied to other geophysical fluids, such as the Ocean. In this context, the work of Bryan (Bryan, 1967) emerges as the first attempt at simulating the ocean state. The first ocean general circulation model with realistic geometries was proposed by Bryan and Cox in 1968 (Bryan and Cox, 1968). This work paved the way for what is today one of the main paradigms in oceanography, the Bryan-Cox-Semtner model (Semtner et al., 1988; Dukowicz et al., 1994; S. M. Griffies, Böning, et al., 2000; Murray et al., 2001). The first attempts at coupling atmosphere and ocean models were undertaken by Manabe et al. (1969), whose work established the paradigm of coupled Earth systems model. Despite this long lasting and continuously growing interest in developing ocean and climate modelling tools, combined with the growing computing capabilities, one fundamental problem remains: the fundamental equations of fluid mechanics present a vast set of interactions that are inherently impossible to fully capture with simulations. A simple mathematical argument is here proposed to describe this problem, and can be summarised with Carl Wunsch's statement "no model has a perfect resolution" (Wunsch, 2015). The equations of motion of a fluid (see Chapter 1 of this thesis) follow a non-linear structure that can be at first studied through a simple advection equation:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{2}$$

If this equation is assigned an analytical initial condition such as $u\left(x, t_0\right) = A\cos\left(kx\right)$, then by Taylor expansion one can infer the state of the system close to the initial state $t = t_0$, that is

$$\begin{aligned} u\left(x,t\right) &= u\left(x,t_0\right) + \left(t - t_0\right)\frac{\partial u\left(x,t_0\right)}{\partial t} + \mathcal{O}\left(\Delta t\right) \\ &= A\cos\left(kx\right) - \left(t - t_0\right)u\frac{\partial u\left(x,t_0\right)}{\partial x} + \mathcal{O}\left(\Delta t\right) \\ &= A\cos\left(kx\right) - \left(t - t_0\right)\frac{A^2k}{2}\sin\left(2kx\right) + \mathcal{O}\left(\Delta t\right). \end{aligned}$$

In the last line is noticeable that the solution has a harmonic frequency $2k$, twice that of the initial condition. A transfer mechanism from large scale to small scale is thus encoded in the nonlinear term, that will create smaller and smaller scales as the motion persists. Similar results hold for two or three dimensional motion (Bailly et al., 2015). This energy

and scale cascade ends at a specific scale, the Kolmogorov dissipation scale (Kolmogorov, 1941), where molecular friction stops the motion, converting kinetic energy into thermal energy through Joule's heating. Any numerical resolution that aims at being fully representative of the flow has to be performed up to the Kolmogorov scale, on the order of 10^{-6}m. Achieving this task for a fluid flow with a given Reynolds number Re requires (in very idealised setting) the capability of handling $N^3 > \mathrm{Re}^{9/4}$ spatial collocation points (Bailly et al., 2015). In addition to this intrinsic limitation on the resolution, another "structural" limitation is present in OGCMs. Numerical resolution of a partial differential equations always starts with the discretization of the unknwown functions, the discretization of the differential operator or a combination of the two. Spatial Galerkin methods are an example of the former approach, where the unknown function is represented as a combination of piecewise continuous functions of some order (Zienkiewicz et al., 2013). Finite differences methods are instead an example of the latter, where the continuous differential operator is replaced by a discrete operator operating on a set of discrete points. The two approaches are often combined, *e.g.* the temporal advance of a spatial Galerkin method with a finite difference scheme in time. Finite differences approximations of continuous differential operators are usually based on Taylor expansion and their order of accuracy depends on the truncation that is adopted. Taking the limit for the discretization step towards zero, the discrete operator should ideally converge to the continuous one. It is a natural conclusion that increasing the resolution of the discrete model should increase the fidelity of the discrete operator. A naive attempt to tackle the problem would be that of hoping in the increase of computational resources, to afford larger and larger computations. This task is however not possible in the next foreseeable future: the smallest time scale in the Ocean can be regarded as $1s$, while the smallest spatial scales can be considered as $10^{-3}m$. A climatic simulation (of duration $\sim 1000y$) of the ocean would require $\sim 10^{27}$ grid cells (S. M. Griffies and Treguier, 2013). Moreover, even if truncation is adopted at a certain scale, such as that given by the Rossby radius of deformation (also called mesoscale), realisation of the importance of sub-mesoscale dynamics for oceanic flows is well documented by a series of recent publications (Capet et al., 2008; Callies et al., 2015; Lévy, Klein, et al., 2010; McWilliams, 2016). These sub-mesoscale eddies, typically of the order of $\mathcal{O}(1)$ to $\mathcal{O}(10)$ kms, have been observed to play a vital role in sustaining the energetic mesoscale eddies which are an order of magnitude larger in scale – $\mathcal{O}(100)$ km. They also play a vital role in characterising large-scale oceanic circulation, tracer mixing in the upper oceans, and in the energy cascade (Capet et al., 2008). Their

presence is observed through satellite field measurements (Durand et al., 2010) and their importance highlighted by numerical studies (Callies et al., 2015; Della Penna et al., 2019). Their contribution to biogeochemical material exchange in the ocean is put forward in the work of Capet et al. (2008) and Lévy, Iovino, et al. (2012). The explicit resolution of these sub-mesoscale eddies in numerical simulation is a computational challenge that requires resolutions smaller than the smallest to-be-resolved eddy (B. Fox-Kemper et al., 2019). This requires grid-sizes smaller than 1km as compared to the 10-100km grid-size generally employed in oceanic general circulation models. It is interesting to note that the explicit resolution of the sub-mesoscale eddies are possible in limited scenarios within which they were shown to have a significant impact on the flow characteristics (Lévy, Klein, et al., 2010; Hurlburt et al., 2000; Siegel et al., 2001). High-resolution modelling requires however finer considerations for the numerical methods employed, posing new challenges to ocean models. The extension of the range of application of this type of model, originally developed for low-resolution large-scale configurations, raises some new challenges of numerical and physical nature (S. M. Griffies and Treguier, 2013) and requires the accurate representation of a wider energy spectrum (Lemarié et al., 2015). One way to address this conundrum is to provided by nesting techniques, where a relatively coarse resolution grid covers the entire domain of computation and is locally refined in areas of interest such as coastlines, straits or areas of known intense activity. This multi-resolution approach is mature enough (Debreu and Blayo, 2008; Debreu, Marchesiello, et al., 2012) to be used to locally reach marginal submesoscales resolving resolutions (Marchesiello et al., 2011). Multiresolution and multigrid strategies are employed to understand the dynamics of passive tracers as well (Jouini et al., 2013; Bricaud et al., 2020). Several general circulation models today implement the well known adaptive grid refinement strategy of AGRIF (Debreu and Blayo, 2002) (see *e.g.* Madec et al., 2019; Auclair et al., 2022). As the resolution increases, a more adequate treatment of the energy transfer between resolved and unresolved scales is required (Sanderson, 1998; Thuburn et al., 2014), while spurious dianeutral mixing remains a challenging aspect in the presence of mesoscale eddies (Ilıcak et al., 2012). Despite the new challenges of a more sophisticated numerics for a high resolution model, coastal models are widespread tools for research. These codes often differ so much from large scale models that their implementation is often either a standalone variant of a general circulation model (see *e.g.* the differences between CROCO and ROMS) or a completely independent project (see *e.g.* Delft3D, MARS Lazure et al., 2008) with possibly completely different approaches (Defina, 2007). The development of these tools

is driven primarily by the necessities of preservation of coastal communities. It is thus not rare to see these models employed in tidal influence studies to assess the possible combined tidal-surge risk (Dumas et al., 2012; Idier et al., 2012) where tides play an important role (*e.g.* in the English Channel) or where storm surge is particularly frequent (*e.g.* the northern Adriatic [2], *cfr* Mel et al., 2013; De Zolt et al., 2006). An attractive alternative option for modelling eddy contributions is through the introduction of stochastic forcing. For instance, Porta Mana et al. (2014) applied a stochastic parametrisation of mesoscale eddies based on resolved variables for a quasi-geostrophic double-gyre configuration with encouraging results. A different stochastic framework for a realistic barotropic model is given by Franzke et al. (2005) through the definition of reduced order models. A semi-analytical, non-linear stochastic closure model was proposed by Grooms and A. J. Majda (2013) under the super parametrisation framework of Grabowski et al. (1999). However, the work of Berloff (2005) suggests that a random forcing does not always account for all fundamentally important non-linear interactions at the resolved scales. In addition, these random forcings are not constrained by physical invariants and thus, may not always be physically relevant. Along this path, two companion methodologies have been introduced by Mémin (2014) and Holm (2015), providing rigorously justified methodology to define stochastic large scales representations of the Navier-Stokes equations (Debussche et al., 2023) conserving energy and circulation, respectively. These two models rely on a stochastic decomposition of the Lagrangian trajectory into a smooth-in-time component induced by the large-scale velocity and a random fast-evolving uncorrelated displacement noise. The solid theoretical background allows the definition of a large-scale representation with a stochastic component representing the subgrid contribution, introducing additional degrees of freedom to be exploited in the modelling of specific phenomena, such as large scale components (Tucciarone et al., 2022; Tucciarone et al., 2023), small-scale turbulence (Harouna et al., 2017; Chandramouli, Mémin, and Heitz, 2020), boundary layer effects (Pinier et al., 2019), convection processes (Jamet et al., 2023) or to devise intermediate models (Harouna et al., 2017; Chapron et al., 2018; Pinier et al., 2019; Bauer et al., 2020a; Cintolesi et al., 2020a). The Location Uncertainty (LU) model (Mémin, 2014) has been applied to the barotropic quasi-geostrophic model (Bauer et al., 2020a; Bauer et al., 2020b), the baroclinic quasi-geostrophic model (Li, Deremble, et al., 2023), the rotating shallow water model (Brecht et al., 2021), the surface quasi-geostrophic (Resseguier,

2. The reader comfortable with engaging with a book in Italian might be interested in "Un giorno, ospite inatteso, arrivò l'alluvione" by Emeritus Professor Luigi D'Alpaos, describing two extreme flooding events in the Northern Adriatic and their impact on the communities.

Li, et al., 2021), hydrostatic primitive equations (Tucciarone et al., 2022; Tucciarone et al., 2023) and recently was formulated for non hydrostatic Boussinesq equations (Tissot, Mémin, et al., 2023), proving its efficacy in structuring the large-scale flow (Bauer et al., 2020a), reproducing long-term statistics (Bauer et al., 2020b) and providing a good trade-off between model error representation and ensemble spread (Resseguier, Li, et al., 2021; Brecht et al., 2021) In this work, the physically consistent stochastic framework of Location Uncertainty (Mémin, 2014; Resseguier, Mémin, and Chapron, 2017a) is applied to an idealised 3D double-gyre configuration using a data-driven model for the noise to characterise unresolved eddy contributions. The double-gyre is an ideal configuration for this study as it shows significant variation in the flow statistics, even in the mean fields, when the horizontal resolution is increased from eddy-resolving to sub-mesoscale resolving (Lévy, Klein, et al., 2010).

Chapter 1

Ocean General Circulation Models

Abstract

This chapter reviews the fundamental equations of fluid mechanics, specialised for the case of geophysical flows. The dynamics and thermodynamics of these flows are not discussed in full details, as it would go beyond the scope of this introduction, instead the aspects connected to the design of the so called Oceanic General Circulation Models (OGCMs) are highlighted. The characteristics of such models go beyond the sole physics necessary to describe the processes of ocean circulation, as they are tied to the method of representing the continuum equation into a discretized set of equations evolving on a computer model. The intent is to present the fundamental equations that are solved in a general circulation model to motivate the exposition of the set of equations presented in Chapter 2.

1.1 Fundamental equations of fluid mechanics

Fluid mechanics, as a branch of continuum mechanics, extends the results of classical point mechanics to parcels of fluid. These parcel have integral properties expressed in terms of a density (*i.e.* the property per unit mass) to be integrated over the volume of the fluid parcels themselves. This volume of fluid, V_t, is not fixed in space and evolves in time. Among the most fundamental results of continuum mechanics, Reynolds' transport theorem (Batchelor, 2000) provides an explicit formula for the temporal variation of the

volume V_t

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{V_t} \mathrm{d}\boldsymbol{x} = \int_{V_t} \nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}, \tag{1.1}$$

with $\boldsymbol{v}$ the vector velocity at which the volume V_t moves, and by extension of the integrated density

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{V_t} \phi\,\mathrm{d}\boldsymbol{x} = \int_{V_t} \frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}, \tag{1.2}$$

with ϕ a general property of the flow per unit mass. It is important to notice that equation (1.1) is the particular case of the second when the function ϕ is chosen to be the indicator function. These two equations constitute the grounding of every balance and conservation statement in fluid mechanics (and continuum mechanics). A conservation statement imposes that the variation in time of the integrated quantity is zero, such as

$$\int_{V_t} \frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} = 0, \tag{1.3}$$

while a balance statement equates the variation in time of a given integrated quantity to the external processes that can modify its temporal evolution, namely a flux through the boundary or a source within the boundary, as

$$\int_{V_t} \frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} = \int_{V_t} \boldsymbol{F}_\phi\,\mathrm{d}\boldsymbol{x}. \tag{1.4}$$

Since the volume V_t can be any generic volume contained in the fluid domain, the localization theorem (Kolmogorov et al., 1975) states that the balance must hold for each point of the domain, hence rendering the local (in an *almost everywhere* sense) conservation statement balance as

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot\boldsymbol{v} = 0, \tag{1.5}$$

and balance statement as

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla\cdot\boldsymbol{v} = \boldsymbol{F}_\phi. \tag{1.6}$$

In the previous equation, the temporal variation is

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} = \frac{\partial\phi}{\partial t} + \boldsymbol{v}\cdot\nabla\phi, \tag{1.7}$$

that is the material derivative, accounting both for local rate of temporal change of the property ϕ, plus a contribution arising from the spatial variation as experienced as the parcel of fluid moves with velocity $\boldsymbol{v}$. The right-hand side is often referred to as the

Eulerian representation of the rate of change, in contrast with the left hand side that is termed Lagrangian representation.

The fundamental statement of fluid mechanics is conservation [1] of mass. Defining mass as $\mathrm{d}M = \rho\,\mathrm{d}V$ with ρ density of the fluid, one finds

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0. \tag{1.8}$$

In cases where the density is not varying in time nor space, the previous equation becomes simply $\nabla \cdot \boldsymbol{u} = 0$, which is a statement about the conservation of the volume (isochoricity) within the flow. This statement could also have been found by defining a conservation statement starting from the evolution of the volume, equation (1.1). Moreover, conservation of mass can be used to derive an important variant of Reynolds' transport theorem for an extensive quantity ϕ

$$\begin{aligned}\frac{\mathrm{D}}{\mathrm{D}t}\int_{V_t} \rho\phi\,\mathrm{d}\boldsymbol{x} &= \int_{V_t} \frac{\mathrm{D}\rho\phi}{\mathrm{D}t} + \rho\phi\nabla\cdot\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} \\ &= \int_{V_t} \rho\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\left[\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla\cdot\boldsymbol{v}\right]\mathrm{d}\boldsymbol{x} = \int_{V_t}\rho\frac{\mathrm{D}\phi}{\mathrm{D}t}\,\mathrm{d}\boldsymbol{x},\end{aligned}$$

where the interior of the square bracket is equal to zero when mass is conserved.

Newton's second law of motion states that the momentum of a body changes under the influence of forces acting on it. For a fluid parcel, these forces can be external (*i.e.* body forces) or internal (*i.e.* surface). The former act through the whole media and in the specific case of the ocean, they are of gravitational and non-inertial nature. The latter contact forces, act on the medium through the boundaries of the parcel itself, and in the context of fluid mechanics these forces are pressure and friction.

The gravitational field is an effective, conservative field that accounts for both the gravitational attraction and centrifugal effects due to Earth's rotation (Batchelor, 2000; Kundu et al., 2015; Vallis, 2017), denoted as $\boldsymbol{\Phi}$. The choice of the form of this gravitational potential, that is often taken as $\boldsymbol{\Phi} \sim gz$ with g constant acceleration of gravity, is important within the aims of tide modelling (Bennett, 2002). Non inertial forces account for the effects of the Earth's rotation, as the observers of the ocean are most likely to be in this rotating frame. The momentum balance is thus written in the non-inertial frame and accounts for the Coriolis acceleration that is felt by the water mass, $\boldsymbol{F}_c = -2\boldsymbol{\Omega} \times \boldsymbol{v}$.

1. Classical mechanics does not involve processes capable of generating or destroying mass, such as in general relativity.

The Earth's rotation vector $\boldsymbol{\Omega}$ is directed South to North and the rotation is anticlockwise when looking towards the North Pole from a higher point of view (with respect to the direction of $\boldsymbol{\Omega}$). Contact forces are given in their most general formulation as

$$\boldsymbol{F}_s = \int_{\partial V} \left[\mathbf{T}\left(r\right) - p\mathbf{I}\right] \cdot \boldsymbol{n}\, \mathrm{d}\varsigma, \tag{1.9}$$

where p is the pressure acting on the boundary ∂V of the fluid parcel V in an isotropic way, while $\mathbf{T}$ is the deviatoric stress tensor, the existence of which is postulated and whose properties (symmetry, trace free, ...) stem from Cauchy's tetrahedron theorem (Landau et al., 1987; Batchelor, 2000; Olivella et al., 2000; Cardin et al., 2013). The choice of the tensor $\mathbf{T}$ is fundamental, as it is often the only component of the equation of motion that accounts for the properties of the fluid considered, as highlighted by the dependence from a *rheologic* property r. Eventually, the most fundamental equation, which is the conservation of linear momentum, can thus be written as

$$\rho \frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} - 2\boldsymbol{\Omega} \times \rho\boldsymbol{v} = -\rho\nabla\boldsymbol{\Phi} + \nabla \cdot \left[\mathbf{T}\left(r\right) - p\mathbf{I}\right], \tag{1.10}$$

where the stress forces $\boldsymbol{F}_s$ were manipulated through Green-Gauss theorem to obtain a volume description. This Euler-Cauchy's equation of motion, becomes the Navier-Stokes equation when the stress tensor reflects the behaviour of a Newtonian fluid

$$\mathbf{T}\left(\mu\right) = -\left(\frac{2}{3}\mu\nabla\cdot\boldsymbol{v}\right)\mathbf{I} + \mu\left[\nabla\boldsymbol{v} + \left(\nabla\boldsymbol{v}\right)^{\mathrm{T}}\right], \tag{1.11}$$

where μ is the dynamic viscosity. The dependence of the momentum tendency on the pressure carries information on the thermodynamics of the fluid, thus requiring the definition of mechanical energy (as the sum of kinetic and potential) and internal energy, building up the notion of total energy. The specific kinetic energy (*i.e.* per unit mass, $K = \boldsymbol{v}^2/2$) budget can be obtained starting from the momentum equation by taking the inner product with the velocity itself, so that it reads

$$\rho\frac{\mathrm{D}K}{\mathrm{D}t} = -\rho\boldsymbol{v}\cdot\nabla\boldsymbol{\Phi} - \boldsymbol{v}\cdot\nabla p + \boldsymbol{v}\cdot\nabla\cdot\mathbf{T}\left(r\right), \tag{1.12}$$

that in his Eulerian form is:

$$\frac{\partial\rho K}{\partial t} + \nabla\cdot\left(\rho\boldsymbol{v}K + \boldsymbol{v}p\right) = -\rho\boldsymbol{v}\cdot\nabla\boldsymbol{\Phi} + p\nabla\cdot\boldsymbol{v} + \boldsymbol{v}\cdot\nabla\cdot\mathbf{T}\left(r\right). \tag{1.13}$$

On the right hand-side there are conversion processes between kinetic energy and other forms of energy: the first term, $-\rho\boldsymbol{v}\cdot\nabla\boldsymbol{\Phi}$ represents the conversion of kinetic energy into potential energy, described by a decreasing kinetic energy when the fluid particles moves to a larger potential state ($\rho\boldsymbol{v}\cdot\nabla\boldsymbol{\Phi}>0$); the second term, $p\nabla\cdot\boldsymbol{v}$ represents the work of pressure forces. The potential energy $\boldsymbol{\Phi}$ budget can be obtained in its general form by definition of temporal variation as

$$\frac{\mathrm{D}\rho\boldsymbol{\Phi}}{\mathrm{D}t}=\boldsymbol{\Phi}\frac{\mathrm{D}\rho}{\mathrm{D}t}+\rho\frac{\mathrm{D}\boldsymbol{\Phi}}{\mathrm{D}t}, \tag{1.14}$$

leading, after application of conservation of mass, to the gravitational potential energy budget equation as

$$\frac{\partial\rho\boldsymbol{\Phi}}{\partial t}+\nabla\cdot(\rho\boldsymbol{\Phi})=\rho\frac{\partial\boldsymbol{\Phi}}{\partial t}+\rho\boldsymbol{v}\cdot\nabla\boldsymbol{\Phi}, \tag{1.15}$$

where the last term $+\rho\boldsymbol{v}\cdot\nabla\boldsymbol{\Phi}$ represents an increasing gravitational potential energy when the fluid particle moves up the gradient and constitutes the most important mean of energy exchange between the two mechanical energy components: kinetic and potential. The first term, the variation in time of gravitational potential energy, is of paramount importance in tidal modelling.

Introducing the mechanical energy as the sum of kinetic K and potential $\boldsymbol{\Phi}$ one has the balance of mechanical energy by summing the previous two results:

$$\frac{\partial}{\partial t}\left[\rho\left(K+\boldsymbol{\Phi}\right)\right]+\nabla\cdot\left[\rho\boldsymbol{v}\left(K+\boldsymbol{\Phi}\right)+\boldsymbol{v}p\right]=\rho\frac{\partial\boldsymbol{\Phi}}{\partial t}+p\nabla\cdot\boldsymbol{v}+\boldsymbol{v}\cdot\nabla\cdot\mathbf{T}\left(r\right) \tag{1.16}$$

The specific total energy E is defined as the sum of the contributions of specific kinetic energy $K=\boldsymbol{v}^2/2$, the specific potential gravitational potential $\boldsymbol{\Phi}$ and the specific internal energy I, that is

$$E=K+\boldsymbol{\Phi}+I, \tag{1.17}$$

with this last embodying the energy of molecular thermal agitation, averaged over the number of molecules to yield a continuous quantity. It is a usual assumption that the energy of a fluid parcel is conserved, meaning that the specific total energy evolves according to the conservation law

$$\frac{\partial\rho E}{\partial t}+\nabla\cdot\mathcal{J}_{E}=\rho\frac{\partial\boldsymbol{\Phi}}{\partial t}, \tag{1.18}$$

for some flux of energy $\mathcal{J}_{E}$. The previous equation is forced by the temporal variation of the geopotential, $\rho\partial_t\boldsymbol{\Phi}$. The flux of energy accounts not only for the aforementioned

fluxes of mechanical energy $\mathcal{J}_{k\boldsymbol{\Phi}} = \rho \boldsymbol{v} \left(K + \boldsymbol{\Phi} \right) + \boldsymbol{v} p$, but also for a heat flux $\mathcal{J}_q$ thermally forcing the system.

The internal energy [2] budget, I, is thus recovered by removing the mechanical energy budget from the conservation of total energy, yielding

$$\rho \frac{\mathrm{D}I}{\mathrm{D}t} = -\nabla \cdot \mathcal{J}_q - p \nabla \cdot \boldsymbol{v} - \boldsymbol{v} \cdot \nabla \cdot \mathbf{T} \left(r \right). \tag{1.19}$$

This equation links the variation of internal energy to the heat fluxes $\mathcal{J}_q$ forcing the system, the work of pressure forces associated with compression/dilatation $-p \nabla \cdot \boldsymbol{v}$ and dissipated by frictional stresses. It constitutes the First principle of thermodynamics for a fluid parcel. Finally, the allowed thermodynamic transformations must be only those respecting the second principle of thermodynamics. Thus, internal energy must be related to entropy, and entropy itself must be related to some physically measurable quantity. Internal energy is thought to be a function of entropy per unit mass ζ, density ρ and salinity S (more generally, chemical potentials), so that $I = I \left(\zeta, \rho, S \right)$ and its infinitesimal variation is

$$\mathrm{d}I = T \, \mathrm{d}\zeta - p \, \mathrm{d}\rho^{-1} + \mu_s \, \mathrm{d}S \tag{1.20}$$

with T is the *in situ* temperature, p pressure and μ_s the relative chemical potential. If the process is quasistatic, the previous equation can be written as a time evolution (in conjunction with mass conservation) as

$$\rho \frac{\mathrm{D}I}{\mathrm{D}t} = \rho T \frac{\mathrm{D}\zeta}{\mathrm{D}t} - p \nabla \cdot \boldsymbol{v} + \mu_s \frac{\mathrm{D}S}{\mathrm{D}t}, \tag{1.21}$$

thus connecting internal energy, entropy, pressure and salinity. The evolution of entropy can be obtained from this equation removing the dependence from internal energy

$$\rho T \frac{\mathrm{D}\zeta}{\mathrm{D}t} = -\nabla \cdot \mathcal{J}_q - \boldsymbol{v} \cdot \nabla \cdot \mathbf{T} \left(r \right) - \rho \mu_s \frac{\mathrm{D}S}{\mathrm{D}t}, \tag{1.22}$$

exposing the evolution of entropy in the fluid by heat forcing, frictional dissipation that increases the heat content of the parcel by Joule's heating and salinity mixing. Moreover, the isoentropic motion of a water parcel is identified as the adiabatic, frictionless and isohaline motion. Since entropy is difficult to measure in practice, an evolution equation for temperature is developed as a useful thermodynamic quantity. Starting from the as-

2. The adjective *specific* has been discarded from this point on, for the sake of notation.

sumption that entropy depends on pressure, temperature and salinity $\zeta = \zeta(p, T, S)$, its infinitesimal variation is

$$\mathrm{d}\zeta = -\frac{\alpha_{_T}}{\rho}\,\mathrm{d}p + \frac{C_{pS}}{T}\,\mathrm{d}T - \frac{\partial \mu_s}{\partial T}\,\mathrm{d}S, \tag{1.23}$$

where classical Maxwell's relations where used to define the thermal expansion coefficient $\alpha_{_T}$, the specific heat at constant pressure and salinity C_{pS} and equivalence between variation of entropy with respect to salinity and variation of chemical potential with respect to temperature. Multiplying by ρT and taking the temporal variation leads to

$$\rho C_{pS}\frac{\mathrm{D}T}{\mathrm{D}t} = \rho T\frac{\mathrm{D}\zeta}{\mathrm{D}t} + T\alpha_{_T}\frac{\mathrm{D}p}{\mathrm{D}t} - \rho T\frac{\partial \mu_s}{\partial T}\frac{\mathrm{D}S}{\mathrm{D}t}. \tag{1.24}$$

Removing the dependence from the entropy one finalizes the temporal evolution of (in situ) temperature:

$$\rho C_{pS}\frac{\mathrm{D}T}{\mathrm{D}t} = T\alpha_{_T}\frac{\mathrm{D}p}{\mathrm{D}t} + \rho\left(\mu_s - T\frac{\partial \mu_s}{\partial T}\right)\frac{\mathrm{D}S}{\mathrm{D}t} - \nabla\cdot\mathcal{J}_q + \boldsymbol{v}\cdot\nabla\cdot\mathbf{T}(r)\,. \tag{1.25}$$

This equation exposes that the temperature of the fluid parcel depends on the adiabatic pressure effects that alters the parcel temperature by expanding or compressing the parcel itself, material changes in the salinity, heat fluxes through the boundaries and frictional dissipation causing Joule's heating.

In addition to this fundamental set of equation, information on the behaviour of the active chemical species involved in the changes of density are needed. They are usually modelled with a transport equation with dissipation. Taking Salinity S as a prototype tracer the equation reads

$$\frac{\partial S}{\partial t} + \nabla\cdot(\rho S\boldsymbol{v} + \rho\mathcal{J}_{_S}) = \rho R_{_S} \tag{1.26}$$

where $\mathcal{J}_{_S}$ summarizes the diffusive fluxes of molecular nature and $R_{_S}$ represents the reaction term, to be interpreted as a source or sink term for the species S. This equation can be applied as well to all the inactive components mixed in the fluid parcel, that do not take an active part in the thermodynamics of the system but for which transport and dispersion is of interest for the modelling of the system. This can include pollutants and contaminants but as well biological species or biologically impacting chemical species.

Thermo-mechanics of a fluid flow

Conservation of mass: (1.27)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0$$

Momentum balance: (1.28)

$$\rho \frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} - 2\boldsymbol{\Omega} \times \rho \boldsymbol{v} = -\rho \nabla \boldsymbol{\Phi} + \nabla \cdot [\mathbf{T}(r) - p\mathbf{I}]$$

Mechanical energy budget: (1.29)

$$\frac{\partial}{\partial t}[\rho (K + \boldsymbol{\Phi})] + \nabla \cdot [\rho \boldsymbol{v} (K + \boldsymbol{\Phi}) + \boldsymbol{v} p] = \rho \frac{\partial \boldsymbol{\Phi}}{\partial t} + p \nabla \cdot \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \cdot \mathbf{T}(r)$$

Internal energy budget (First law of thermodynamics): (1.30)

$$\rho \frac{\mathrm{D}I}{\mathrm{D}t} = -\nabla \cdot \mathcal{J}_q - p \nabla \cdot \boldsymbol{v} - \boldsymbol{v} \cdot \nabla \cdot \mathbf{T}(r)$$

Conservation of total energy: (1.31)

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot [\rho \boldsymbol{v} (K + \boldsymbol{\Phi}) + \boldsymbol{v} p + \mathcal{J}_q] = \rho \frac{\partial \boldsymbol{\Phi}}{\partial t}$$

Entropy budget (Second law of thermodynamics): (1.32)

$$\rho T \frac{\mathrm{D}\zeta}{\mathrm{D}t} = -\nabla \cdot \mathcal{J}_q - \boldsymbol{v} \cdot \nabla \cdot \mathbf{T}(r) - \rho \mu_s \frac{\mathrm{D}S}{\mathrm{D}t}$$

Temperature budget: (1.33)

$$\rho C_{pS} \frac{\mathrm{D}T}{\mathrm{D}t} = T \alpha_{_T} \frac{\mathrm{D}p}{\mathrm{D}t} + \rho \left(\mu_s - T \frac{\partial \mu_s}{\partial T} \right) \frac{\mathrm{D}S}{\mathrm{D}t} - \nabla \cdot \mathcal{J}_q + \boldsymbol{v} \cdot \nabla \cdot \mathbf{T}(r)$$

Temperature budget: (1.34)

$$\frac{\partial S}{\partial t} + \nabla \cdot (\rho S \boldsymbol{v} + \rho \mathcal{J}_{_S}) = \rho R_{_S}$$

Equation of state: (1.35)

$$\rho = \rho (T, S, p) .$$

1.2 Oceanic General Circulation Models

So far the equation that has been presented are approximated only in a minimal sense: the continuum hypothesis is employed and can be viewed as an approximation of the

discrete matter, the rotation of the Earth is assumed constant and the geometry is assumed spherical without considering deviations of the geoid from the sphere. These equations describe the full motion of a thermodynamically active fluid, with temporal scales ranging from seconds to millennia and spatial scales from the molecular scale (smaller than a millimetre) to the domain scale. As stated in the introduction of this thesis, this range is far from approachable with modern day technology or technology in the foreseeable future. The processes of interest must thus be defined, targeted, and a various techniques to filter out the non-relevant processes must be adopted. The oceanic general circulation is defined (see McWilliams, 1996) as the currents on horizontal space scale and time scale greater than the mesoscale, which for the ocean is of the order of 100km and three months, with the associated fields of pressure, density, temperature and salinity, plus all the other elements needed to establish balances between these fields. Ocean modelling (Baylor Fox-Kemper et al., 2019) is a branch of numerical modeling that focuses on the representation of the physical mechanisms governing the evolution of ocean physical properties, namely T, S, u, v, and w, where T is the temperature, S the salinity, and u, v, and w, the horizontal and vertical components of the velocity $\boldsymbol{u}$. The pioneering work of Bryan and Cox (1968) led to the first global solutions obtained by Cox (1975), establishing the fundamental paradigm for General Circulation Models of the ocean.

The inviscid Navier-Stokes equation, or Euler equation, is obtained from the Navier-Stokes equation by neglecting the effects of molecular viscosity, thus setting $\mathbf{T}(\mu) = 0$, that is

$$\rho \frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} - 2\boldsymbol{\Omega} \times \rho \boldsymbol{v} = -\nabla p - \rho \nabla \boldsymbol{\Phi}. \tag{1.36}$$

This latter constitutes the principal model for the motion of a fluid for which viscous effects can be neglected, that is the case of the Ocean and the Atmosphere. This first important simplification is motivated by experimental results (Vallis, 2017), as the magnitude of dynamical viscosity of water or air is orders of magnitude smaller than other players in the equation (such as advection, Coriolis acceleration and pressure gradients). The fundamental fluid dynamics of the Ocean circulation is thus described by the Euler equation on the rotating Earth for a compressible fluid, namely seawater, which is composed of water plus a suite of dissolved salts that occur in nearly constant ratio but variable amount (Teramoto, 1992), with an empirically determined equation of state (McDougall et al., 2011). The dominant forces in the ocean are associated with pressure, Coriolis force and gravity. The thermodynamic of oceanic water is simplified in such a way that it only accounts for two tracers, which are temperature and salinity, affected

by circulation, internal sources and sinks but no molecular mixing. These two tracers are assumed to be conservative (S. Griffies et al., 2008).

Euler equations on the sphere

Horizontal momentum: (1.37a,b)

$$\frac{\mathrm{D}u}{\mathrm{D}t} - \left(2\Omega + \frac{u}{r\cos\theta}\right)(v\sin\theta - w\cos\theta) = -\frac{1}{\rho r\cos\theta}\frac{\partial p}{\partial\lambda}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + \frac{wv}{r} + \left(2\Omega + \frac{u}{r\cos\theta}\right)u\sin\theta = -\frac{1}{\rho r}\frac{\partial p}{\partial\theta}$$

Vertical momentum: (1.38)

$$\frac{\mathrm{D}w}{\mathrm{D}t} - 2\Omega u\cos\theta - \frac{u^2+v^2}{r} = \frac{1}{\rho}\frac{\partial p}{\partial r} - g$$

Temperature and salinity: (1.39a,b)

$$\frac{\partial T}{\partial t} + \frac{1}{r\cos\theta}\frac{\partial(uT)}{\partial\lambda} + \frac{1}{r\cos\theta}\frac{\partial}{\partial\theta}(vT\cos\theta) + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 wT\right) = R_{_T}$$

$$\frac{\partial S}{\partial t} + \frac{1}{r\cos\theta}\frac{\partial(uS)}{\partial\lambda} + \frac{1}{r\cos\theta}\frac{\partial}{\partial\theta}(vS\cos\theta) + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 wS\right) = R_{_S}$$

Mass conservation: (1.40)

$$\frac{\partial\rho}{\partial t} + \frac{1}{r\cos\theta}\frac{\partial(u\rho)}{\partial\lambda} + \frac{1}{r\cos\theta}\frac{\partial}{\partial\theta}(v\rho\cos\theta) + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 w\rho\right) = 0$$

Equation of state: (1.41)

$$\rho = \rho(T, S, p)\,.$$

The material derivative in spherical coordinates (and flux form) is given by

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} = \frac{\partial\phi}{\partial t} + \frac{1}{r\cos\theta}\frac{\partial(u\phi)}{\partial\lambda} + \frac{1}{r\cos\theta}\frac{\partial}{\partial\theta}(v\phi\cos\theta) + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 w\phi\right). \tag{1.42}$$

A general circulation model derives from these equations with several further simplifications. When considering large scale dynamics, the ocean horizontal length scale is much larger than the vertical length scale [3]. Quoting Prof. Robert Higdon: *The ocean is shallow,*

3. The ISO 216 international standard paper size A4 has dimension of L_{297} =297mm height times L_{210} =210mm width, with a thickness of H =0.065mm (standard 80g/m^2), with aspect ratios L_{297}/H ∼4570 and L_{210}/H ∼3230. The Ocean, with its average depth H =3.7km has an equator-wise aspect ratio of L_{eq}/H =40075km/3.7km∼10831 and a North-South aspect ratio

your coffee mug is deep. This consideration about the smallness of the aspect ration H/L is known as the *shallow water condition*. It has important consequences and it is accurate in the case of small Rossby and Froude numbers (the former is the ratio between nonlinear advection and rotation, the latter is a measure of the importance of stratification). For this reason it is useful to introduce a horizontal-vertical splitting notation. Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the chosen reference and let's define $\nabla_{\mathrm{H}} = [\partial_{\mathbf{e}_1}, \partial_{\mathbf{e}_2}, 0]$ so that $\nabla = \nabla_{\mathrm{H}} + \partial_z \mathbf{e}_3$. This splitting holds for Cartesian references but holds as well for spherical coordinate, and it will be used extensively in the following. The first fundamental consequence is that the vertical acceleration can be neglected and thus the ocean can be regarded as hydrostatic. From volume conservation one can infer that $W \sim \frac{H}{L}U$, thus the vertical velocity has a much smaller magnitude than the horizontal velocity, $w \ll u$. Moreover, the vertical momentum equation can be scaled as

$$\frac{W}{T} + \frac{UW}{H} + \frac{W^2}{H} + \Omega U \sim \left| \frac{1}{\rho} \frac{\partial p}{\partial r} \right| - g \tag{1.43}$$

where all the first three terms scale as $(H/L)(U/L)$ and thus they are small with respect to all the other terms. The Coriolis term can be considered small as well, leading to hydrostatic balance: pressure vertical gradients and gravity are in balance. Furthermore, the *shallow fluid approximation* can be taken, defining $r = a + z$ with $a \gg z$ distance to the centre of the earth, that can be assumed constant, hence replacing r with a and $\partial/\partial r$ with $\partial/\partial z$. Finally, the *traditional approximation* is used to neglect all non-horizontal components of the Coriolis term. The resulting equations are so–called *Primitive Equations*. The term "primitive" refers to the fact that the momentum equations are expressed in terms of velocity, or momentum, in contrast with other simplified models that use prognostic equations for derived quantities such as vorticity, streamfunction and divergence (S. Griffies et al., 2008). These equations are summarised below.

L_{NS}/H =20000km/3.7km∼5405. Roughly speaking, printing out the map of the ocean on an A4 sheet of paper will render a picture of the ocean with a thickness twice that the original.

Primitive Equations

Horizontal momentum: (1.44a,b)

$$\frac{\mathrm{D}u}{\mathrm{D}t} - 2\Omega v \sin\theta + \frac{uv\tan\theta}{a} = -\frac{1}{\rho r\cos\theta}\frac{\partial p}{\partial \lambda}$$
$$\frac{\mathrm{D}v}{\mathrm{D}t} + 2\Omega u \sin\theta + \frac{u^2\tan\theta}{a} = -\frac{1}{\rho r}\frac{\partial p}{\partial \theta}$$

Vertical momentum: (1.45)

$$\frac{\partial p}{\partial z} = \frac{g\rho}{\rho_0}$$

Temperature and salinity: (1.46a,b)

$$\frac{\partial T}{\partial t} + \frac{1}{a\cos\theta}\frac{\partial (uT)}{\partial \lambda} + \frac{1}{a\cos\theta}\frac{\partial}{\partial \theta}(vT\cos\theta) + \frac{\partial}{\partial z}(wT) = 0$$
$$\frac{\partial S}{\partial t} + \frac{1}{a\cos\theta}\frac{\partial (uS)}{\partial \lambda} + \frac{1}{a\cos\theta}\frac{\partial}{\partial \theta}(vS\cos\theta) + \frac{\partial}{\partial z}(wS) = 0$$

Incompressibility: (1.47)

$$\frac{1}{a\cos\theta}\frac{\partial (u)}{\partial \lambda} + \frac{1}{a\cos\theta}\frac{\partial}{\partial \theta}(v\cos\theta) + \frac{\partial}{\partial z}(w) = 0$$

Equation of state: (1.48)

$$\rho = \rho(T, S, z)\,.$$

1.3 Density variation and Boussinesq approximation

Under conditions typically satisfied by the Ocean or the Atmosphere (see Spiegel et al., 1960 for a review), the Boussinesq hypothesis (Boussinesq, 1903; Zeytounian, 2003; Vallis, 2017) suggests that the importance of density variations can be neglected everywhere except when it directly causes buoyant forces, that means acting in conjunction with gravity g. The fundamental statement is that in the conservation of mass (1.27) written in material form (that is, using (1.7) instead of the conservative form) the relative variations of density $\rho^{-1}\frac{\mathrm{D}\rho}{\mathrm{D}t}$ are much smaller than the effects of the converging or diverging flow velocity $\nabla\cdot\boldsymbol{v}$. This stems from a decomposition of the density

$$\rho(\boldsymbol{x}, t) = \rho_0 + \delta\rho(\boldsymbol{x}, t)\,, \tag{1.49}$$

where the deviation $\delta\rho\left(\boldsymbol{x},t\right)$ is supposed to be much smaller than the reference density, *i.e.* $\delta\rho \ll \rho_0$. In the Ocean, variations of the density of seawater are of the order of 1% (Gray et al., 1976), thus validating this assumption. The mass conservation equation is replaced by volume conservation, with consequent exclusion of acoustic waves from the model, consistent with the low Mach number of the ocean. In the momentum equations, the reference density ρ_0 can be used in all terms, with exception to the gravity term where the ratio between the fluctuation and the reference density is used, introducing thus the buoyancy

$$b = g\frac{\rho\left(\boldsymbol{x},t\right)-\rho_0}{\rho_0}. \tag{1.50}$$

Despite support arbitrary non-linear equations of state in a fully energetically and thermodynamically consistent manner, while retaining the inherent simplicity of a divergent-free velocity (Tailleux, 2010), simple Boussinesq thermodynamics can be defined (as shown in Kundu et al., 2015 and Vallis, 2017) as

$$\rho C_p \frac{\mathrm{D}\Theta}{\mathrm{D}t} = -\nabla\cdot\mathcal{J}_q, \tag{1.51}$$

where Θ is some conservative temperature. Since active tracers modify the density of the fluid, the fluid dynamics is influenced by density gradients through the hydrostatic relation, with the equation of state providing the fundamental bridge between dynamics and thermodynamics. The Equation of Seawater (EoS) is an empirical non-linear thermodynamic relationship linking seawater density, ρ, to a number of state variables, most typically temperature, salinity and pressure. The full description of the EoS is quite complex (McDougall et al., 2011) and its implementation in general circulation models is challenging (Roquet et al., 2015). A simplified EoS inspired by Vallis (2017) is often chosen, and it provides simplistic linear representation of many processes, such as cabbeling and thermobaricity (Roquet et al., 2015). With such an equation of state there is no longer a distinction between conservative and potential temperature, as well as between absolute and practical salinity. S-EoS takes the following expression:

$$\begin{aligned} d_a\left(T,S,Z\right) = \frac{1}{\rho_0}\left[-a_0\left(1+0.5\lambda_1 T_a+\mu_1 z\right)\cdot T_a + b_0\left(1-0.5\lambda_2 S_a+\mu_2 z\right)\cdot S_a\right. \\ \left. -\nu T_a S_a + \beta_p\left(p-p_0\right)\right] \end{aligned}$$

with $T_a = T - 10$, $S_a = S - 35$, $\rho_0 = 1026\text{Kg/m}^3$ and where d_a is the thermodynamic potential obtained by differentiation of the Gibbs function of Seawater with respect to pressure (see Vallis, 2017, equation 1.146). Various approximation of EoS can be specified simply by changing the associated coefficients by changing the thermobaric coefficients (μ_1, μ_2) or by changing the cabbeling coefficients $(\lambda_1, \lambda_2, \nu)$. Keeping non-zero value to a_0 and b_0 provides a simple and linear EoS function of T and S:

$$\rho = \rho_0 \left[1 - \beta_T \left(T - T_0\right) + \beta_S \left(S - S_0\right) + \beta_p \left(p - p_0\right)\right], \tag{1.52}$$

with $\beta_T \approx 2 \times 10^{-4}\text{K}^{-1}$, $\beta_S \approx 10^{-3}\text{g/kg}$ and $\beta_p \approx 4.4 \times 10^{-10}\text{m s}^{-2}\text{kg}^{-1}$. This linear equation retails the three main drivers of density variation: compression of seawater by the action of pressure, thermal expansion due to increase of temperature and, if salinity increase, haline contraction.

1.4 Sub-Grid Scales modelling (SGS)

The primitive equations are typically run with resolution that can spam from $\sim$ 100km (very large scale) to $\sim$ 1km (fine scale). In both cases, several processes are left behind in the model, while the ocean dynamics is affected by the cumulative effect of turbulent processes occurring at all scales. For these reason, truncation of the representable scales must be performed for practical simulation, and the inclusion of the effects of the unresolved scales is of paramount importance. Standard practice is to approach the task with a Reynolds Averaged Navier Stokes (Reynolds, 1883; Batchelor, 2000) approach, by substituting the model variables u, v, w, T, S with their mean components plus the fluctuation, that is

$$\boldsymbol{u} = \overline{\boldsymbol{u}} + \boldsymbol{u}', \quad T = \overline{T} + T', \quad S = \overline{S} + S'. \tag{1.53}$$

Nonlinearity of the advection term introduces a turbulent stress tensor in each prognostic equation

$$\text{SGS}^{\boldsymbol{u}} = \nabla \cdot \overline{u_j' u_i'}, \quad \text{SGS}^{T} = \nabla \cdot \overline{u_j' T'}, \quad \text{SGS}^{S} = \nabla \cdot \overline{u_j' S'}. \tag{1.54}$$

The choice for the representation of the eddy quantities, i.e $\overline{u_j' u_i'}$, $\overline{u_j' T'}$ and $\overline{u_j' S'}$, is a crucial part of the modeller task, and usually means parametrizing these turbulent effects without resolving them. Among all the debated choices for this closure term, that must be specified by the modeller, the simplest one is Prandtl's K-closure model (also referred to as

Boussinesq eddy viscosity model, *cfr* Oertel, 2010), that prescribes direct proportionality between the eddy quantity and the shear:

$$\overline{u_j' u_i'} = A_{ij}^m \frac{\partial u_i}{\partial x_j}, \quad \overline{u_j' T'} = A_{ij}^T \frac{\partial T}{\partial x_j}, \quad \overline{u_j' S'} = A_{ij}^S \frac{\partial S}{\partial x_j}. \tag{1.55}$$

On the scales of interest for large scale modelling of the Ocean, the mixing occurs along directions that are approximately horizontal, decorrelated from the vertical direction. If isotropic turbulence is assumed in the horizontal, usual forms of the sub-grid scale models are

$$\begin{aligned}
\mathrm{SGS}^{\boldsymbol{u}} &= A^m \Delta \boldsymbol{u} + \frac{\partial}{\partial z}\left(A_{33}^m \frac{\partial \boldsymbol{u}}{\partial z} \right), \\
\mathrm{SGS}^{T} &= A^T \Delta \boldsymbol{u} + \frac{\partial}{\partial z}\left(A_{33}^T \frac{\partial T}{\partial z} \right), \\
\mathrm{SGS}^{S} &= A^S \Delta \boldsymbol{u} + \frac{\partial}{\partial z}\left(A_{33}^S \frac{\partial S}{\partial z} \right).
\end{aligned}$$

Empirical observations indicate that in the interior of the ocean, the rate of transport along principal directions of mixing can be as much as 10^8 times the rate of transport in the orthogonal direction. This phenomenon was linked to direction of neutral buoyancy, that are direction lying to the tangent plane of the isopycnal surfaces. If a parcel of water moves across this surface, then the density contrast with the surrounding environment will act as a restoring force toward its original position (assuming that molecular mixing is not acting at this scale). On the other hand, any adiabatic motion that occur along neutral surfaces encounters no such impediment, resulting thus in a preferred direction of motion. This is often expressed by introducing a diffusion tensor that is locally aligned to the isopycnal surfaces (Redi, 1982)

$$\mathbf{K}_{\mathrm{R}} = \frac{A_{_H}}{\rho_x^2 + \rho_y^2 + \rho_z^2} \begin{bmatrix} \rho_x^2 + \rho_y^2 + \epsilon\rho_z^2 & (1-\epsilon)\,\rho_x\rho_y & (1-\epsilon)\,\rho_x\rho_z \\ (1-\epsilon)\,\rho_x\rho_y & \rho_x^2 + \rho_y^2 + \epsilon\rho_z^2 & (1-\epsilon)\,\rho_y\rho_z \\ (1-\epsilon)\,\rho_x\rho_z & (1-\epsilon)\,\rho_y\rho_z & \rho_x^2 + \rho_y^2 + \epsilon\rho_z^2 \end{bmatrix} \tag{1.56}$$

where $A_{_H}$ is the horizontal diffusion coefficient, the smallness parameter ϵ is defined as $A_{_z} = \epsilon A_{_H}$ and the notation $\rho_i = \partial_i \rho$ for each axis was introduced for notational convenience. In the same spirit, irreversible mixing processes are usually modelled with a downgradient diffusion that tends to flatten isopycnal surfaces thus transferring the

energy from potential to kinetic (P. R. Gent et al., 1990; Peter R. Gent, 2011).

These observations hold for the interior of the ocean. The surface layer present exclusive characteristics and it is usually termed *mixed layer* to describe this characteristic. The surface of the ocean is the most energy-dense turbulent zone, with turbulence driven by wind, waves and buoyancy forcing, resulting at the end in creation of potential energy through mixing and entrainment, viscous dissipation or energy loss through internal waves.

1.5 Boundary conditions

Equations (1.44a,b) to (1.47) are valid in the interior of the ocean, that means for $\eta_b(x,y) \leq z \leq \eta(x,y,t)$, where η_b is the bottom coordinate and η is the surface elevation. At these two boundaries, exchanges with the solid earth or with the atmosphere happen and must be prescribed. The momentum equations have dynamic boundary conditions, relating the momentum stress at the surface with the wind stress, that is

$$A_{33}^m \frac{\partial u_i}{\partial z}\bigg|_{z=\eta} = \frac{1}{\rho}\tau_j^w. \tag{1.57}$$

In a similar fashion, at the bottom the momentum stress must be equal to the bottom friction:

$$A_{33}^m \frac{\partial u_i}{\partial z}\bigg|_{z=\eta_b} = f(u_i), \tag{1.58}$$

where the functional for of f depends from the specific friction parametrization. Similarly, temperature is forced at the surface as

$$A_{ij}^T \frac{\partial T}{\partial x_j}\bigg|_{z=\eta} = \frac{Q_s + Q_l + Q_b}{C_p \rho_0}, \tag{1.59}$$

with the sensible heat flux Q_s (due to the air-sea temperature difference), the latent heat flux Q_l (due to evaporation), and the long-wave back radiation Q_b. Note that upward (positive) fluxes of temperature lead to cooling at the surface. For rigid lid models, salinity has the same structure with the effects of evaporation E and precipitation P on the vertical diffusive flux

$$A_{ij}^S \frac{\partial S}{\partial x_j}\bigg|_{z=\eta} = -(\mathrm{E} - \mathrm{P})\, S. \tag{1.60}$$

Bottom forcing for temperature can be taken into account for long-term climatic simulations.

1.6 Surface pressure and external gravity waves

Total pressure at a given depth z is composed of a surface pressure p_s at a reference geopotential surface (that is $z = 0$) and a hydrostatic pressure p_h, in such a way that

$$p\left(x, y, z, t\right) = p_s\left(x, y, t\right) + p_h\left(x, y, z, t\right), \tag{1.61}$$

with the hydrostatic term computed as

$$p_h\left(x, y, z, t\right) = \int_{\varsigma=z}^{\varsigma=0} g\rho\left(T, S, \varsigma\right) \, \mathrm{d}\varsigma. \tag{1.62}$$

The surface pressure can be computed in different ways. The first method, historically referred to as *rigid-lid* approximation, consists in keeping the surface of the ocean still at a geopotential reference and solving an equation for pressure. As introduced in Bryan and Cox (1968), it sets the time tendency of the sea surface to zero, thus implying a non-divergent barotropic velocity and so allowing the construction of a streamfunction for the vertically integrated velocity. If on the one hand this approximation facilitates efficient time stepping by introducing an elliptic equation, on the other hand it introduces an elliptic equation which is difficult to solve in complex geometries and removes the possibility to model tides or direct freshwater forcing (Huang, 1993). The second method, employed in almost all current general circulation models, consist in integrating a prognostic equation for the sea surface elevation, namely

$$\frac{\partial \eta}{\partial t} = -\nabla \cdot \left[\left(\eta - \eta_b\right) \overline{\boldsymbol{U}}\right] + P - E, \tag{1.63}$$

describing the evolution of external gravity waves, propagating at a speed of roughly $\sqrt{gH}$. This last equation is not enough to describe the dynamic of external gravity waves (EGW) and a prognostic equation for $\overline{\boldsymbol{U}}$ must be provided as well. To prescribe this additional dynamical equation the linear stratified primitive equation can be projected onto a set of orthogonal normal modes, under several assumptions, highlighting a fast barotropic mode and a slower baroclinic mode (Demange et al., 2019). A classical time-stepping method consists then to introduce a splitting procedure to advance the barotropic and

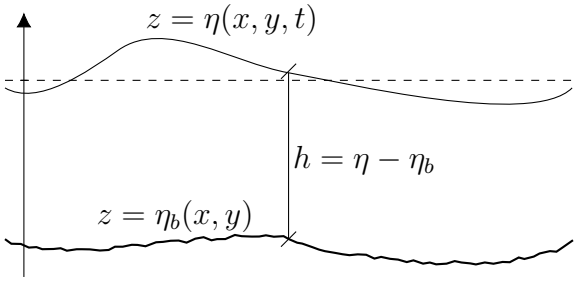


Figure 1.1 – Single layer shallow water model. The bottom topography is a time independent spatially varying surface designated with $z = \eta_b(x, y)$, while the free surface is a material surface defined as $z = \eta(x, y, t)$. The total thickness of the layer is simply $h = \eta - \eta_b$. The level $z = 0$ is intended as the rest position of the free surface, and is represented with the horizontal dashed line.

the baroclinic parts differently. A usual assumption, detailed in Killworth et al. (1991) consists in assuming that the external mode is vertically constant. Following Demange et al. (2019) the system of equations describing the depth-averaged motion is:

$$\begin{aligned} \frac{\partial \overline{\boldsymbol{U}}}{\partial t} + g\nabla\eta &= -\frac{1}{\rho_0}\nabla\overline{p_h} + G\left(\boldsymbol{u}\right) \\ \frac{\partial \eta}{\partial t} &= -\nabla\cdot\left[\left(\eta - \eta_b\right)\overline{\boldsymbol{U}}\right], \end{aligned}$$

where $G\left(\boldsymbol{u}\right)$ is the vertical average of the advective, diffusive terms of the baroclinic momentum equations, and $\overline{p_h}$ is the vertical average of the internal pressure gradient. The right hand side of the momentum equations is usually kept constant during the short time-step time evolution, while the integral of p_h depends on the sea surface elevation η itself, making the convergence of this integral problematic. In the following, the depth-integrated equations of motion are derived following Blumberg et al. (1987) and Killworth et al. (1991) to show the terms taken into consideration. A Barotropic-Baroclinic splitting is defined as

$$u = \frac{1}{\eta - \eta_b}\int_{\eta_b}^{\eta} u\,\mathrm{d}\zeta + u', \qquad v = \frac{1}{\eta - \eta_b}\int_{\eta_b}^{\eta} v\,\mathrm{d}\zeta + v', \tag{1.64}$$

where the vertically integrated velocities (from bottom to top) will be denoted with overlined capital letters $\overline{U}$ and $\overline{V}$ and denoted barotropic velocities, while the residual components u' and v' will be denoted as baroclinic velocities.

The first step is integrating from bottom to top the continuity equation

$$\nabla_{\text{H}} \cdot \boldsymbol{u}_{\text{H}} + \frac{\partial w}{\partial z} = 0 \tag{1.65}$$

using the kinematic boundary conditions

$$\begin{aligned} w|_{\eta} &= \frac{\partial \eta}{\partial t} + \boldsymbol{u}_{\text{H}}\left(\boldsymbol{x}_{\text{H}}, \eta, t\right) \cdot \nabla_{\text{H}} \eta \\ w|_{\eta_b} &= \frac{\partial \eta_b}{\partial t} + \boldsymbol{u}_{\text{H}}\left(\boldsymbol{x}_{\text{H}}, \eta_b, t\right) \cdot \nabla_{\text{H}} \eta_b, \end{aligned}$$

that leads to

$$\int_{\eta_b}^{\eta} \nabla_{\text{H}} \cdot \boldsymbol{u}_{\text{H}} \, \mathrm{d}\zeta + \frac{\partial \eta}{\partial t} + \boldsymbol{u}\left(\boldsymbol{x}_{\text{H}}, \eta, t\right) \cdot \nabla_{\text{H}} \eta - \frac{\partial \eta_b}{\partial t} - \boldsymbol{u}\left(\boldsymbol{x}_{\text{H}}, \eta_b, t\right) \cdot \nabla_{\text{H}} \eta_b = 0.$$

Using Leibnitz integration rule

$$\begin{aligned} \nabla_{\text{H}} \cdot \int_{\eta_b}^{\eta} \boldsymbol{u}\left(\boldsymbol{x}_{\text{H}}, z, t\right) \, \mathrm{d}\zeta = \boldsymbol{u}\left(\boldsymbol{x}_{\text{H}}, \eta, t\right) \cdot \nabla_{\text{H}} \eta - \boldsymbol{u}\left(\boldsymbol{x}_{\text{H}}, \eta_b, t\right) \cdot \nabla_{\text{H}} \eta_b \\ + \int_{\eta_b}^{\eta} \nabla_{\text{H}} \cdot \boldsymbol{u}\left(\boldsymbol{x}_{\text{H}}, z, t\right) \, \mathrm{d}\zeta \end{aligned}$$

one finally arrives to

$$\frac{\partial \eta}{\partial t} + \nabla \cdot \left[\left(\eta - \eta_b\right) \overline{\boldsymbol{U}}\right] = 0. \tag{1.66}$$

This equation is the familiar linear gravity wave equation and provides a prognostic equation for η. Integrating from bottom to top the momentum equation

$$\frac{\partial u_i}{\partial t} + \boldsymbol{u} \cdot \nabla_{\text{H}} u_i + w \frac{\partial u_i}{\partial z} + \epsilon_{jki} f_j u_k = -\frac{\partial}{\partial x_i} \frac{p'}{\rho_0} + F^i \tag{1.67}$$

one obtains

$$\begin{aligned} \int_{\eta_b}^{\eta} \frac{\partial u_i}{\partial t} \, \mathrm{d}\zeta + \int_{\eta_b}^{\eta} \nabla_{\text{H}} \cdot \left(\boldsymbol{u} u_i\right) \, \mathrm{d}\zeta + \int_{\eta_b}^{\eta} \frac{\partial w u_i}{\partial z} \, \mathrm{d}\zeta \\ = -\epsilon_{jki} f_j \int_{\eta_b}^{\eta} u_k \, \mathrm{d}\zeta + \int_{\eta_b}^{\eta} \left[-\frac{\partial}{\partial x_i} \frac{p'}{\rho_0} + F^i\right] \, \mathrm{d}\zeta. \end{aligned}$$

The terms in the left hand side can be modified using Leibniz integral rule, with temporal

tendency reading

$$\int_{\eta_b}^{\eta} \frac{\partial}{\partial t}\boldsymbol{u}\left(\boldsymbol{x}_{\mathrm{H}}, z, t\right)\,\mathrm{d}\zeta = \frac{\partial}{\partial t}\int_{\eta_b}^{\eta}\boldsymbol{u}\left(\boldsymbol{x}_{\mathrm{H}}, z, t\right)\,\mathrm{d}\zeta - \boldsymbol{u}\left(\boldsymbol{x}_{\mathrm{H}}, \eta, t\right)\cdot\nabla_{\mathrm{H}}\eta + \boldsymbol{u}\left(\boldsymbol{x}_{\mathrm{H}}, \eta_b, t\right)\cdot\nabla_{\mathrm{H}}\eta_b,$$

and momentum transport reading

$$\begin{aligned}\int_{\eta_b}^{\eta}\nabla_{\mathrm{H}}\cdot\left(\boldsymbol{u}u_i\right)\,\mathrm{d}\zeta &= \nabla_{\mathrm{H}}\cdot\int_{\eta_b}^{\eta}\boldsymbol{u}u_i\,\mathrm{d}\zeta - \left(\boldsymbol{u}u_i\right)\cdot\nabla_{\mathrm{H}}\eta + \left(\boldsymbol{u}u_i\right)\cdot\nabla_{\mathrm{H}}\eta_b\\ \int_{\eta_b}^{\eta}\frac{\partial w u_i}{\partial z}\,\mathrm{d}\zeta &= w\left(\eta\right)u_i\left(\eta\right) + w\left(\eta_b\right)u_i\left(\eta_b\right)\\ &= u_i\left(\eta\right)\left[\frac{\partial\eta}{\partial t} + \boldsymbol{u}\left(\boldsymbol{x}_{\mathrm{H}}, \eta, t\right)\cdot\nabla_{\mathrm{H}}\eta\right] - u_i\left(\eta_b\right)\left[\boldsymbol{u}\left(\boldsymbol{x}_{\mathrm{H}}, \eta_b, t\right)\cdot\nabla_{\mathrm{H}}\eta_b\right].\end{aligned}$$

Summing them one finds

$$\frac{\partial\overline{U}_i}{\partial t} + \nabla_{\mathrm{H}}\cdot\int_{\eta_b}^{\eta}\boldsymbol{u}u_i\,\mathrm{d}\zeta + \epsilon_{jki}f_j\overline{U}_k = \int_{\eta_b}^{\eta}\left[-\frac{\partial}{\partial x_i}\frac{p'}{\rho_0} + F^i\right]\,\mathrm{d}\zeta. \tag{1.68}$$

To complete the system, one adds and subtract a shallow water-like pressure gradient $g\nabla\eta$, so the resulting equations are

$$\begin{aligned}\frac{\partial\overline{U}_i}{\partial t} + \epsilon_{jki}f_j\overline{U}_k + g\nabla\eta &= \mathcal{F}\left(p', \eta\right)\\ \frac{\partial\eta}{\partial t} + \nabla\cdot\left[\left(\eta - \eta_b\right)\overline{\boldsymbol{U}}\right] &= 0,\end{aligned}$$

with

$$\mathcal{F}\left(p', \eta\right) = g\nabla\eta + \nabla_{\mathrm{H}}\cdot\int_{\eta_b}^{\eta}\boldsymbol{u}u_i\,\mathrm{d}\zeta + \int_{\eta_b}^{\eta}\left[-\frac{\partial}{\partial x_i}\frac{p'}{\rho_0} + F^i\right]\,\mathrm{d}\zeta. \tag{1.69}$$

These equations associated with the fast barotropic mode are usually handled with a sub-stepping procedure (Shchepetkin et al., 2005; Shchepetkin et al., 2009), where an internal procedure solves this set of equation with a time step much smaller than that of the full model. The term $\mathcal{F}\left(p', \eta\right)$ is a coupling operator between the baroclinic and the barotropic mode, it contains the slow dynamics of the model and it is used as a constant forcing during the sub-cycling. It is important to notice that this last term contains the divergence of the momentum transport, rendering the barotropic equations linear in their momentum. This choice is common in several ocean models (*e.g.* Madec et al., 2019), but not strictly necessary as several other models implement non-linear barotropic equations (*e.g.* Auclair et al., 2022). For this reason, in the next chapter a stochastic variant of the

shallow water equations is introduced following the same procedure.

Conclusions

This chapter provided a brief and overly simplified description of geophysical fluid mechanics equations and related modelling. The complexity of the subject is such that for each one of the topics here introduced a vast literature is present, and many other were not introduced. The choice of the topics here introduced reflects the main challenges that a modeller has to tackle when embarking into the development of a stochastic general circulation model.

Chapter 2

Location uncertainty principle

Abstract

This chapter aims to introduce the stochastic framework of Location Uncertainty. The presentation is detailed while being free of stochastic calculus as much as possible, with only physical arguments and procedures that are well established in the fluid mechanics practice. The interested reader is however invited to refer to the last section of this chapter to find the mathematical proof of every important formula.

2.1 Location uncertainty

In the Location Uncertainty (LU) formalism, the Lagrangian displacement $\mathbf{X}_t$ associated to a fluid particle is described by a stochastic differential equation of the type:

$$\mathrm{d}\mathrm{X}_t^i(\boldsymbol{x}_0) = v^i(\mathbf{X}_t, t)\,\mathrm{d}t + \sigma_t \mathrm{d}\mathrm{B}_t^i(\mathbf{X}_t), \quad i = 1, 2, 3 \tag{2.1}$$

where $\mathbf{X}_t \colon \Omega \times \mathbb{R}^+ \to \Omega$ is the fluid flow map, that is the trajectory followed by a fluid particle starting at point $\mathbf{X}|_{t=0}(\boldsymbol{x}) = \boldsymbol{x}_0$ of the bounded domain Ω and the index $i = 1, \ldots, 3$ indicates the spatial coordinates x, y, z. The first component, $\boldsymbol{v}$ in vector form, represents the smooth, resolved velocity field of the flow. This term is the response of the equations of motions, solved on a grid of a given resolution, and so it is supposed to be both spatially and temporally correlated. Its integral in time is assumed to be of bounded variation. The second term, that we will indicate as $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t$ in vector form, is a stochastic process that assembles the unresolved component, turbulent effects and the uncertainties

of the flow. This stochastic contribution is referred to as *noise* and defined as

$$\sigma_t \mathrm{d}\mathrm{B}_t^i \left(\mathbf{X}_t\right) = \int_\Omega \breve{\sigma}_{ik} \left(\mathbf{X}_t, \boldsymbol{y}, t\right) \, \mathrm{d}\mathrm{B}_t^k \left(\boldsymbol{y}\right) \mathrm{d}\boldsymbol{y}, \quad i = 1, 2, 3 \tag{2.2}$$

that is the application of an integration kernel $\breve{\sigma}$ to an I_3 cylindrical Wiener process $\mathbf{B}_t$, combination of independent one-dimensional cylindrical Wiener processes, not differentiable in time. In the previous equation, Einstein's summation notation was adopted. The result of such operation is a centred Gaussian process with two-points, two-times covariance tensor defined as

$$\begin{aligned} \boldsymbol{Q} \left(\boldsymbol{x}, \boldsymbol{y}, t, s\right) &= \mathbb{E} \left[\left(\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \left(\boldsymbol{x}\right)\right) \left(\boldsymbol{\sigma}_s \mathrm{d}\mathbf{B}_s \left(\boldsymbol{y}\right)\right)^{\mathrm{T}} \right] \\ &= \delta \left(t - s\right) \delta \left(\boldsymbol{x} - \boldsymbol{y}\right) \mathrm{d}t \int_\Omega \breve{\sigma}_{ik} \left(\boldsymbol{x}, \boldsymbol{z}, t\right) \breve{\sigma}_{kj} \left(\boldsymbol{y}, \boldsymbol{z}, s\right) \, \mathrm{d}\boldsymbol{z}, \end{aligned}$$

meaning that $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \sim \mathcal{N} \left(0, \boldsymbol{Q}\right)$. Further details on the properties of the noise are given in chapter 3. The strength of the noise is measured by its one-point, one-time covariance, denoted as $\boldsymbol{a}$, given by the diagonal components of the covariance tensor per unit of time, that is

$$\boldsymbol{a} \left(\boldsymbol{x}, t\right) := \frac{\boldsymbol{Q} \left(\boldsymbol{x}, \boldsymbol{x}, t, t\right)}{\mathrm{d}t} = \boldsymbol{\sigma}_t \left(\boldsymbol{x}\right) \boldsymbol{\sigma}_t^{\mathrm{T}} \left(\boldsymbol{x}\right), \tag{2.3}$$

leading to the explicit relation:

$$a_{ij} \left(\boldsymbol{x}, t\right) = \int_\Omega \breve{\sigma}_{ik} \left(\boldsymbol{x}, \boldsymbol{y}, t\right) \breve{\sigma}_{kj} \left(\boldsymbol{x}, \boldsymbol{y}, t\right) \, \mathrm{d}\boldsymbol{y}. \tag{2.4}$$

From the relation $\boldsymbol{a} \left(\boldsymbol{x}, t\right) = \boldsymbol{\sigma}_t \left(\boldsymbol{x}\right) \boldsymbol{\sigma}_t^{\mathrm{T}} \left(\boldsymbol{x}\right)$ it is important to recognize that the variance tensor is symmetric and positive semi-definite at any point $\boldsymbol{x}$ of the domain, property that will be employed in the construction of the stochastic Reynolds transport theorem. From this construction it emerges the fact that the displacement is an *Itô diffusion* process, and in the following it will be intensively used in its vector form:

$$\mathrm{d}\mathbf{X}_t \left(\mathbf{X}_0\right) = \boldsymbol{v} \left(\mathbf{X}_t, t\right) \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \left(\mathbf{X}_t\right). \tag{2.5}$$

It is important to state that the expressions contained in (2.1) and (2.5) are shortcut notations for the integral relations:

$$\begin{aligned} \mathrm{X}_t^i = \mathrm{X}_0^i + \int_0^t v^i\left(\mathbf{X}_s, s\right)\,\mathrm{d}s + \int_0^t \int_\Omega \breve{\sigma}_{ik}\left(\mathbf{X}_s, \boldsymbol{y}, t\right)\mathrm{dB}_s^k\left(\boldsymbol{y}\right)\mathrm{d}\boldsymbol{y}, \quad i = 1,2,3, \\ \mathbf{X}_t = \mathbf{X}_0 + \int_0^t \boldsymbol{v}\left(\mathbf{X}_s, s\right)\,\mathrm{d}s + \int_0^t \boldsymbol{\sigma}_s\left(\mathbf{X}_s\right)\,\mathrm{d}\mathbf{B}_s, \end{aligned} \tag{2.6}$$

with $\mathbf{X}_0$ initial condition for the flow. These equations constitute the real mathematical description of the stochastic process as a *semi-martingale* (Kunita, 1997; Da Prato et al., 2014; Brémaud, 2020), defined for each point of $\boldsymbol{x} \in \Omega$ and each time $t \in T \subseteq \mathbb{R}^+$. A total velocity can then be formally defined as

$$\begin{aligned} \boldsymbol{V}\left(\boldsymbol{x}, t\right) = \boldsymbol{v}\left(\boldsymbol{x}, t\right) + \boldsymbol{\sigma}_t \dot{\mathbf{B}}_t\left(\boldsymbol{x}\right), \\ V^i\left(\boldsymbol{x}, t\right) = v^i\left(\boldsymbol{x}, t\right) + \sigma_t \dot{\mathrm{B}}_t^i\left(\boldsymbol{x}\right), \quad i = 1,2,3 \end{aligned} \tag{2.7}$$

where the uncertainty term $\mathrm{d}\mathbf{B}_t$ is not differentiable in time in the classical sense and so its derivative $\dot{\mathbf{B}}_t$ must be thought in terms of weak derivatives, that is

$$\int_T \varphi\left(t\right)\dot{\mathbf{B}}_t\left(\boldsymbol{x}\right)\,\mathrm{d}t = -\int_T \varphi'\left(t\right)\mathbf{B}_t\left(\boldsymbol{x}\right)\,\mathrm{d}t \qquad \forall \varphi \in C_0^\infty\left(T\right). \tag{2.8}$$

From the physical point of view, $\boldsymbol{v}$ is a smooth-in-time random velocity, while $\boldsymbol{\sigma}_t\dot{\mathbf{B}}_t$ represents the noise: a generalized stochastic process that has to be defined in the space of tempered distributions. As introduced, the noise term $\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t$ is defined as centred, that means $\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t \sim \mathcal{N}\left(0, \boldsymbol{Q}\right)$, with $\mathbf{B}$ an I_3-Wiener process defined on a filtered probability space $\{\Omega, \mathcal{F}, \mathrm{P}, (\mathcal{F}_t)_t\}$ and where $(\mathcal{F}_t)_t$ is the filtration associated with $\mathbf{B}_t$. However, the following argument opens to a useful generalisation of non-centred noises, used extensively in Chapter 3. Let $\left(\mathcal{F}_t\right)_t$ be the filtration associated with $\left(\mathbf{B}_t\right)_t$. A non centred Wiener process shifted by a random process $\left(\mathbf{Y}_t\right)_t$ can be defined as:

$$\tilde{\mathbf{B}}_t = \mathbf{B}_t + \int_0^t \mathbf{Y}_s\,\mathrm{d}s. \tag{2.9}$$

Under good properties of $\left(\mathbf{Y}_t\right)_t$ ($\mathcal{F}_t$-measurability, almost sure L^2-integrability and Novikov condition) there exists a measure Q such that:

— $(\tilde{\mathbf{B}}_t)_t$ is a $\mathrm{Q}-$Wiener process;
— the Radon-Nikodym derivative (which is well defined due to Novikov condition

(Brémaud, 2020; Da Prato et al., 2014) writes:

$$\frac{\mathrm{d}\,\mathrm{Q}}{\mathrm{d}\,\mathrm{P}} = \exp\left[-\frac{1}{2}\left(\int_0^t \mathbf{Y}_s\,\mathrm{d}\mathbf{B}_s + \int_0^t \mathbf{Y}_s^2\,\mathrm{d}s\right)\right]. \tag{2.10}$$

With the non centred (under P) random process $\tilde{\mathbf{B}}_t$, we can rewrite the equations with respect to $\tilde{\mathbf{B}}_t$ as

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = \boldsymbol{\sigma}_t \mathrm{d}\tilde{\mathbf{B}}_t - \boldsymbol{\sigma}_t \mathrm{Y}_t\,\mathrm{d}t. \tag{2.11}$$

Denoting $\boldsymbol{\sigma}_t \mathrm{Y}_t$ as $\boldsymbol{\mu}_t$ one can write the Lagrangian displacement under LU as

$$\mathrm{d}\mathbf{X}_t = \left[\boldsymbol{v}\left(\mathbf{X}_t, t\right) - \boldsymbol{\mu}_t\left(\mathbf{X}_t\right)\right]\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\tilde{\mathbf{B}}_t\left(\mathbf{X}_t\right). \tag{2.12}$$

Under Q the Wiener process $\mathrm{d}\tilde{\mathbf{B}}_t$ is centred thus the writing of $\mathrm{d}\mathbf{X}_t$ has the same form as (2.5) but under a new measure. All the arguments provided in the following will hold for this process under Q.

2.2 Stochastic transport theorem

This section highlights the effects of the scale separation introduced in the previous section. In particular, the Stochastic Reynolds Transport Theorem (SRTT), that provides an Eulerian representation to the Lagrangian flow described, is introduced without proof (that is provided in Section 2.9). With respect to the classical Reynolds Transport Theorem, the additional complexity is inherited by the one-dimensional cylindrical Wiener processes $\mathrm{B}_t^k\left(\boldsymbol{x}\right)$, used to separate the scales of motion, that is not differentiable in time, and so it requires special care when left evolving in time. Let's consider a semi-martingale $\theta\left(\boldsymbol{x}, t\right)$ of the type

$$\theta\left(\boldsymbol{x}, t\right) = \theta\left(\boldsymbol{x}, 0\right) + \int_0^t g\left(\boldsymbol{x}, s\right)\,\mathrm{d}s + \int_0^t \int_\Omega f_k\left(\boldsymbol{x}, \boldsymbol{y}, s\right)\,\mathrm{d}\mathrm{B}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y} \tag{2.13}$$

where g, f are processes that, when integrated in time, are of bounded variation. The process $\theta\left(\boldsymbol{x}, t\right)$ is supposed to be regular, that means to have bounded spatial gradients and to be twice differentiable in space. Moreover, the process $\theta\left(\boldsymbol{x}, t\right)$, when transported

by a flow as (2.6), that means the mapping $\boldsymbol{x} \to \mathbf{X}_t$, has a total differential of the form

$$\begin{aligned} \mathrm{d}\theta\left(\mathbf{X}_t, t\right) = \mathrm{d}_t\theta\left(\mathbf{X}_t, t\right) \quad &+ \frac{\partial\theta\left(\mathbf{X}_t, t\right)}{\partial x_i}\mathrm{dX}_t^i + \frac{1}{2}a_{ij}\left(\mathbf{X}_t, t\right)\frac{\partial^2\theta\left(\mathbf{X}_t, t\right)}{\partial x_i \partial x_j}\,\mathrm{d}t \\ &+ \int_\Omega \breve{\sigma}_{ij}\left(\mathbf{X}_t, \boldsymbol{y}, t\right)\frac{\partial}{\partial x_i}f_j\left(\mathbf{X}_t, \boldsymbol{y}, t\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t, \end{aligned} \tag{2.14}$$

where $\mathrm{d}_t\theta = \theta\left(\mathbf{X}_t, t + \mathrm{d}t\right) - \theta\left(\mathbf{X}_t, t\right)$ is the temporal increment at a fixed location, while a_{ij} is the variance tensor introduced in (2.4). This statement is the result of the application of *Itô-Wentzell formula* (see Kunita, 1997, theorem 3.8) applied to the semi-martingale $\theta\left(\boldsymbol{x}, t\right)$, represented in Equation (2.13), composed with the Lagrangian stochastic flow $\mathbf{X}_t$ of Equation (2.6). The last formula has inside the term $f_j\left(\mathbf{X}_t, \boldsymbol{y}, t\right)$ which is still not related to any physical process of interest. To establish this relation, a study of the process $\theta\left(\mathbf{x}, t\right)$ while transported by an Itô diffusion process as (2.6) and conserved, that means $\mathrm{d}\theta\left(\mathbf{X}_t, t\right) = 0$, has to be performed. In this case, the increment in time of $\theta\left(\mathbf{x}, t\right)$ takes the form of

$$\begin{aligned} \mathrm{d}_t\theta \quad = -\Bigg[\Bigg(v^i - \frac{1}{2}\frac{\partial a_{ki}}{\partial x_k} \quad &+ \int_\Omega \breve{\sigma}_{ji}\frac{\partial\breve{\sigma}_{kj}}{\partial x_k}\,\mathrm{d}\boldsymbol{y}\Bigg)\,\mathrm{d}t \\ &+ \int_\Omega \breve{\sigma}_{ik}\mathrm{dB}_t^k\,\mathrm{d}\boldsymbol{y}\Bigg]\frac{\partial\theta}{\partial x_i} + \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial\theta}{\partial x_k}\right)\mathrm{d}t, \end{aligned} \tag{2.15}$$

that can be written in vector notation as

$$\mathrm{d}_t\theta = -\left[\left(\boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}}\left(\nabla\cdot\boldsymbol{\sigma}_t\right)\right)\mathrm{d}t + \boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\right]\cdot\nabla\theta + \frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\theta\right)\mathrm{d}t, \tag{2.16}$$

with the introduction of the notation $\boldsymbol{v}_s = \frac{1}{2}\nabla\cdot\boldsymbol{a}$. All the unknown components of the semi-martingale (2.13) are now directly linked to the motion of the fluid, as this formula expresses the time variation along a fluid particle trajectory. Finally, the main results of Mémin (2014) can be presented.

Theorem 1 (Stochastic Reynolds transport theorem)**.** *Consider a physical quantity $\theta\left(\boldsymbol{x}, t\right)$ within a material volume $V_t \subset \mathbb{R}^3$, transported by a stochastic flow as* (2.6) *and such that it can be written in the semi-martingale form*

$$\theta\left(\boldsymbol{x}, t\right) = \theta\left(\boldsymbol{x}, 0\right) + \int_0^t g\left(\boldsymbol{x}, s\right)\,\mathrm{d}s + \int_0^t\int_\Omega f_k\left(\boldsymbol{x}, \boldsymbol{y}, s\right)\,\mathrm{dB}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}s, \tag{2.17}$$

where f, g are locally bounded processes. Then, the stochastic Reynolds transport theorem

(SRTT) reads:

$$\begin{aligned}\mathrm{d}\int_{V_t}\theta\,\mathrm{d}\boldsymbol{x} \;=\int_{V_t}\Bigg\{\mathrm{d}_t\theta+\frac{\partial}{\partial x_k}\left[\theta\left(u_k-\frac{1}{2}\frac{\partial a_{ik}}{\partial x_i}\right)\mathrm{d}t+\theta\int_\Omega\breve{\sigma}_{kj}\mathrm{dB}_t^j\,\mathrm{d}\boldsymbol{y}\right]\\ -\frac{1}{2}\int_{V_t}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial\theta}{\partial x_k}\right)\mathrm{d}t\Bigg\}\,\mathrm{d}\boldsymbol{x},\end{aligned}\tag{2.18}$$

or in vector formulation as:

$$\mathrm{d}\int_{V_t}\theta\,\mathrm{d}\boldsymbol{x}=\int_{V_t}\left\{\mathrm{d}_t\theta+\nabla\cdot\left[\theta\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\theta\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\theta\right)\,\mathrm{d}t\right\}\,\mathrm{d}\boldsymbol{x}.\tag{2.19}$$

This theorem lays the foundations of the LU formalism, and it has been used extensively since its appearance in Mémin (2014) to develop a family of models of various physical phenomena. The assumption behind the derivation of these models is that the forcing applied to the system is of bounded variation. An extension to stochastic forcing can be defined, as examined and formalised rigorously in Resseguier, Mémin, and Chapron (2017a) and Tissot, Mémin, et al. (2023), where the final result is presented as follows.

Theorem 2 (Extended stochastic Reynolds transport theorem)**.** *Within the assumptions of Theorem 1 and in the presence of a stochastic balance such as*

$$\mathrm{d}\int_{V_t}\theta\,\mathrm{d}\boldsymbol{x}=\int_{V_t}\left(\Theta_t\,\mathrm{d}t+\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t\right)\,\mathrm{d}\boldsymbol{x},\tag{2.20}$$

the stochastic Reynolds' transport theorem reads

$$\begin{aligned}\int_{V_t}\mathrm{d}_t\theta+\nabla\cdot\left[\theta\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\theta\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\theta\right)\,\mathrm{d}t+\nabla\cdot\left(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma\right)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\ =\int_{V_t}\Theta_t\,\mathrm{d}t+\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Dropping the volume integral, one has

$$\mathrm{d}_t\theta+\nabla\cdot\left[\theta\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\theta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\theta\right)\,\mathrm{d}t+\nabla\cdot\left(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma\right)\,\mathrm{d}t=\Theta_t\,\mathrm{d}t+\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t.\tag{2.21}$$

In the previous equation, the stochastic component of the forcing is defined as

$$\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t=\int_\Omega\breve{\Theta}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},t\right)\,\mathrm{dB}_t^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y},\tag{2.22}$$

where is of paramount importance to state that the Brownian term defining this stochastic forcing is the same that drives the motion. The full proof of this theorem is given in section 2.9. It is worth noticing that the term $\nabla \cdot (\boldsymbol{\sigma}_t \boldsymbol{\Theta}_\sigma)$ is just a compact notation for

$$\nabla \cdot (\boldsymbol{\sigma}_t \boldsymbol{\Theta}_\sigma) = \frac{\partial}{\partial x_j} \int_\Omega \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t\right) \breve{\Theta}_\sigma^k \left(\mathbf{X}_t, \boldsymbol{y}, t\right) \, \mathrm{d}\boldsymbol{y}, \tag{2.23}$$

and it must be evaluated differently for each observable it is applied to. A remarkable example of application of this theorem can be found in Tissot, Mémin, et al. (2023).

2.3 Stochastic transport operator

In the previous sections the characteristics of a transported semi-martingale were discussed. It is worth noticing that these characteristics can be encoded in a spatio-temporal operator $\mathbb{D}_t$ that can be linked to the usual deterministic material derivative.

Definition 1 (Stochastic transport operator)**.** If $\theta\left(\mathbf{x}, t\right)$ is a semi-martingale of type (2.13), that is transported by a flow as (2.6) and is conserved, then the stochastic material derivative is defined as:

$$\mathbb{D}_t\theta = \mathrm{d}_t\theta + \left[\left(v_k - \frac{1}{2}\frac{\partial a_{ik}}{\partial x_i} + \int_\Omega \breve{\sigma}_{jk} \frac{\partial \breve{\sigma}_{ij}}{\partial x_i} \, \mathrm{d}\boldsymbol{y}\right) \mathrm{d}t \, + \int_\Omega \breve{\sigma}_{kj} \mathrm{dB}_t^j \left(\boldsymbol{y}\right) \, \mathrm{d}\boldsymbol{y}\right] \frac{\partial \theta}{\partial x_k} - \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial \theta}{\partial x_k}\right) \mathrm{d}t,$$

in vector notation as

$$\mathbb{D}_t\theta = \mathrm{d}_t\theta + \left[\left(\boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma}_t\right)\right) \mathrm{d}t + \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t\right] \cdot \nabla\theta - \frac{1}{2}\nabla \cdot \left(\boldsymbol{a}\nabla\theta\right) \, \mathrm{d}t,$$

The terms appearing in this operator can be physically interpreted. The first term is the *increment in time* at a fixed location of the process θ, that is $\mathrm{d}\theta = \theta\left(\mathbf{X}_t, t + \Delta t\right) - \theta\left(\mathbf{X}_t, t\right)$. This contribution plays the role of the partial time derivative for a process that is not time differentiable. The term enclosed in the square brackets is a *stochastic advection displacement*, and is induced by a slowly varying, time correlated process named *modified drift* velocity and defined as

$$\boldsymbol{v}^\star = \boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma}_t\right) \tag{2.24}$$

and by a fast evolving, time uncorrelated noise $\boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t$. This term involves the multiplication of the noise with the solution gradient, that is itself dependent on the noise: it is thus an advective generalized *multiplicative noise*, therefore non Gaussian. This type of noise is often denoted as *transport noise* (Flandoli et al., 2023). The component $\boldsymbol{v}_s = \frac{1}{2}\nabla \cdot \boldsymbol{a}$ is a statistically induced velocity that in the following will be referred to as *Itô-Stokes drift* as it can be shown that an analogy exists between this term and the Stokes drift of deterministic interaction between waves and the mean flows (Bauer et al., 2020a; Li, 2021). It is worth noticing that $\nabla \cdot \mathbf{a}$ is dimensionally consistent with a velocity, as $\boldsymbol{a} \sim L^2/T$. The variance tensor $\boldsymbol{a}$, which is positive definite, can be interpreted as a diffusion tensor. Indeed, it can be observed that the last term in the transport operator is a divergence of a gradient, that is $\nabla \cdot (\boldsymbol{a}\nabla\theta)$, that is exactly the form of a diffusion term in the standard deterministic PDE setting. Giving a less heuristic explanation, this term can be seen to be a dissipation term under suitable boundary conditions, that means

$$\int_\Omega \theta \nabla \cdot (\boldsymbol{a}\nabla\theta) \, \mathrm{d}\boldsymbol{x} = -\int_\Omega \nabla\theta^{\mathrm{T}} \boldsymbol{a} \nabla\theta \, d\boldsymbol{x} \leq 0 \tag{2.25}$$

after application of Gauss theorem and noticing that $\boldsymbol{a}$ is positive semi-definite. In the end, this mechanism is physically identified as an eddy diffusion due to the unresolved scales, the greater the noise variance, the more intense the mixing. The stochastic advective velocity is important because it represents the transport velocity for a stochastic process. It is important to outline that this eddy diffusion is not artificially introduced but comes rigorously from the scale separation performed in (2.1). Following Resseguier, Mémin, and Chapron (2017a) one can consider the transport of a characteristic function, i.e. using (2.19) with $\theta = 1$, to introduce an evolution equation for the Jacobian determinant J of the flow:

$$\mathbb{D}_t J - J\nabla \cdot [\boldsymbol{v}^\star \, \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t] = 0. \tag{2.26}$$

The Jacobian of the flow represents the change in the integral measure. The previous equation provides then a condition for a stochastic flow to not change the integral measure during the flow, that is to be isochoric, and which yields

$$\nabla \cdot [\boldsymbol{v}^\star \, \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t] = 0. \tag{2.27}$$

For a transported scalar and isochoric flow, it can be demonstrated that the transport operator conserves energy (Resseguier, Mémin, and Chapron, 2017a).

2.3.1 Relation with the material derivative

A relation between the classical material derivative and the transport operator can be set from theorem 1, enabling a better interpretation of what has been presented so far. In the deterministic setting the *material* or *substantial derivative* is usually defined as the variation in time of a given property F along time following a specific particle (material point, identified with its trajectory $\boldsymbol{r}$) of the continuous medium. Introducing the notation D_t for the material derivative one can write

$$\mathrm{D}_t F\left[\boldsymbol{r}\left(\boldsymbol{r}_0, t_0; t\right), t\right] = \frac{\mathrm{d}}{\mathrm{d}t} F\left[\boldsymbol{r}\left(\boldsymbol{r}_0, t_0; t\right), t\right]. \tag{2.28}$$

In particular, applying the chain rule one has

$$\mathrm{D}_t F\left[\boldsymbol{r}\left(\boldsymbol{r}_0, t_0; t\right), t\right] = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial r_1}\frac{\mathrm{d}r_1}{\mathrm{d}t} + \frac{\partial F}{\partial r_2}\frac{\mathrm{d}r_2}{\mathrm{d}t} + \frac{\partial F}{\partial r_3}\frac{\mathrm{d}r_3}{\mathrm{d}t}. \tag{2.29}$$

If $\boldsymbol{r}$ and $\boldsymbol{x}$ define the same point in space, that is $\boldsymbol{x} = \boldsymbol{r}\left(\boldsymbol{r}_0, t_0; t\right)$, then the classical bridge between the Lagrangian and the Eulerian representation is found as

$$\mathrm{D}_t F\left(\boldsymbol{x}, t\right) = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial x_1} u + \frac{\partial F}{\partial x_2} v + \frac{\partial F}{\partial x_3} w. \tag{2.30}$$

The quantity $\mathrm{D}_t F$ is the total differential in time of F, in the stochastic case the resulting formula is functionally different but the meaning remains the same. Equation (2.14) defines the total differential in time of a random process θ along the stochastic flow $\mathbf{X}_t$ as

$$\begin{aligned}\mathrm{d}\theta\left(\mathbf{X}_t, t\right) &= \mathrm{d}_t\theta + \left[\left(\boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}}\left(\nabla\cdot\boldsymbol{\sigma}_t\right)\right)\mathrm{d}t + \boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\right]\cdot\nabla\theta - \frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\theta\right)\mathrm{d}t \\ &:= \mathbb{D}_t\theta\end{aligned}$$

This equation is valid when there is no stochastic forcing inducing changes in the transport of θ. Conversely, when the transport is forced by a stochastic process one finds that the total differential in time is given by

$$\begin{aligned}\mathrm{d}\theta\left(\mathbf{X}_t, t\right) &= \mathrm{d}_t\theta + \nabla\cdot\left[\theta\left(\boldsymbol{v} - \boldsymbol{v}_s\right)\mathrm{d}t + \theta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right] - \frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\theta\right)\mathrm{d}t + \nabla\cdot\left(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma\right)\mathrm{d}t \\ &= \mathbb{D}_t\theta + \mathrm{Tr}\left(\boldsymbol{\sigma}_t^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma\right)\mathrm{d}t.\end{aligned}$$

The last term, $\mathrm{Tr}\left(\boldsymbol{\sigma}_t^{\mathrm{T}} \nabla \boldsymbol{\Theta}_\sigma\right)$ is present whenever there is a correlation between the stochastic forcing and the stochastic transport, *i.e.* if the two terms depend on the same Brownian motion. In the case the forcing depends on a Brownian motion that is independent from the Brownian motion of the transport, the correlation cancels. Finally, one can state

$$\mathrm{D}_t\theta = \mathrm{d}\theta = \mathbb{D}_t\theta + \mathrm{Tr}\left(\boldsymbol{\sigma}_t^{\mathrm{T}} \nabla \boldsymbol{\Theta}_\sigma\right)\, \mathrm{d}t. \tag{2.31}$$

2.3.2 Distributivity of the stochastic transport operator

The distributivty of the operator $\mathbb{D}_t$ is a fundamental inquiry to proceed deriving conservation statements and balance equations. Considering two variables f and g, the evolution of which is given by

$$\begin{aligned}\mathbb{D}_t f &= F_t\, \mathrm{d}t + \boldsymbol{F}_\sigma \cdot \mathrm{d}\mathbf{B}_t \\ \mathbb{D}_t g &= G_t\, \mathrm{d}t + \boldsymbol{G}_\sigma \cdot \mathrm{d}\mathbf{B}_t\end{aligned}$$

the evolution of the product of f and g is then

$$\mathbb{D}_t\left(fg\right) = g\mathbb{D}_t f + f\mathbb{D}_t g + \boldsymbol{F}_\sigma \cdot \boldsymbol{G}_\sigma\, \mathrm{d}t - \boldsymbol{\sigma}_t \boldsymbol{F}_\sigma \cdot \nabla g\, \mathrm{d}t - \boldsymbol{\sigma}_t \boldsymbol{G}_\sigma \cdot \nabla f\, \mathrm{d}t, \tag{2.32}$$

with proof and detailed description of the right hand side given in section 2.9. It is important to notice that the classical product rule holds when $\boldsymbol{F}_\sigma = \boldsymbol{G}_\sigma = 0$.

2.3.3 Non-dimensional transport operator

Following a classical procedure, the stochastic Navier-Stokes equations are now written in dimensionless form. Introducing the *characteristic scales* L and U, respectively for the length and the velocity of the fluid motion, T, a typical time scale of the motion (that can be expressed in terms of other scales), one can write each variable as

$$\boldsymbol{x} = L\hat{\boldsymbol{x}}, \qquad \boldsymbol{v} = U\hat{\boldsymbol{v}}, \qquad t = T\hat{t}. \tag{2.33}$$

Similarly for these widely used normalizations (Batchelor, 2000), in this stochastic settings one needs to define characteristic scales also for the stochastic variables. From Equation

(2.2) it is clear that $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t$ must be dimensionally consistent with $\mathrm{d}\mathbf{X}_t$, thus

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = L_\sigma \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}_t, \tag{2.34}$$

with L_σ characteristic scale of the stochastic displacement. The Brownian motion by itself scales as $\mathbf{B}_t \sim \sqrt{T_\sigma}\hat{\mathbf{B}}_t$, where T_σ is the temporal scale of the stochastic motion. These two scales are *a priori* not the same scales as the deterministic one introduced in (2.33). Introducing A as the characteristic scale for the variance tensor $\boldsymbol{a} = A\hat{\boldsymbol{a}}$, one can then define the scaling of the kernel as $\boldsymbol{\sigma}_t = \sqrt{A}\hat{\boldsymbol{\sigma}}_t$ by definition in comparison with (2.4). The scaling of the variance tensor (and thus of the kernel $\boldsymbol{\sigma}$) can be related recalling the definition of noise

$$\sigma \mathrm{dB}_t^i(\boldsymbol{x}) = \int_\Omega \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{y}, t)\, \mathrm{dB}_t^k(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}, \tag{2.35}$$

so that both $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = L_\sigma \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}_t$ and $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = \sqrt{AT_\sigma}\hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}_t$ must hold, providing $A \sim \frac{L_\sigma^2}{T_\sigma}$. Finally, the stochastic scaling reads

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = L_\sigma \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}_t, \qquad \boldsymbol{\sigma}_t = \frac{L_\sigma}{\sqrt{T_\sigma}}\hat{\boldsymbol{\sigma}}_t, \qquad \mathbf{B}_t \sim \sqrt{T_\sigma}\hat{\mathbf{B}}_t, \qquad \boldsymbol{a} \sim \frac{L_\sigma^2}{T_\sigma}\hat{\boldsymbol{a}}. \tag{2.36}$$

To simplify the calculations that are going to be carried out in the next sections, the non-dimensional form of the stochastic transport operator and of the impulse are introduced separately for typographical convenience. The procedure shown is applied only to the variables defining the transport, not to the quantity that is actually transported. It is important to remark that in the previous equation θ is kept with its dimensions. If θ scales as $\theta = Q\hat{\theta}$ one can write $\hat{\mathbb{D}}_t\theta = \Theta\hat{\mathbb{D}}_t\hat{\theta}$. This will be helpful in simplifying the non-dimensional Navier-Stokes equations. Substituting these variables in the transport operator one finds:

$$\begin{aligned}\hat{\mathbb{D}}_t\theta = \mathrm{d}_t\theta + \Bigg\{ \Big[\frac{UT}{L}\hat{\boldsymbol{v}} - \frac{AT}{L^2}\frac{1}{2}\hat{\nabla}\cdot\hat{\boldsymbol{a}} + \frac{AT}{L^2}\hat{\boldsymbol{\sigma}}_t^{\mathrm{T}}\left(\hat{\nabla}\cdot\hat{\boldsymbol{\sigma}}_t\right)\Big]\,\mathrm{d}\hat{t} \\ + \frac{L_\sigma}{L}\hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}_t \Bigg\}\cdot\hat{\nabla}\theta - \frac{AT}{L^2}\frac{1}{2}\hat{\nabla}\cdot\left(\hat{\boldsymbol{a}}\hat{\nabla}\theta\right)\mathrm{d}\hat{t}.\end{aligned}$$

In fluid mechanics a usual choice for the temporal scale is the so called *advective time scale*, defined as $T = L/U$, measuring the time needed to transport a particle with the characteristic velocity of the fluid for the characteristic length scale. With this choice one

has

$$\hat{\mathbb{D}}_t\theta = \mathrm{d}_t\theta + \left\{\left[\hat{\boldsymbol{v}} - \frac{AT}{L^2}\frac{1}{2}\hat{\nabla}\cdot\hat{\boldsymbol{a}} + \frac{AT}{L^2}\hat{\boldsymbol{\sigma}}_t^{\mathrm{T}}\left(\hat{\nabla}\cdot\hat{\boldsymbol{\sigma}}_t\right)\right]\mathrm{d}\hat{t} + \frac{L_\sigma}{L}\hat{\boldsymbol{\sigma}}_t\mathrm{d}\hat{\mathbf{B}}_t\right\}\cdot\hat{\nabla}\theta - \frac{AT}{L^2}\frac{1}{2}\hat{\nabla}\cdot\left(\hat{\boldsymbol{a}}\hat{\nabla}\theta\right)\mathrm{d}\hat{t}.$$

In this novel framework a non-dimensional parameter is introduced,

$$\Upsilon = \frac{UL}{A} = \frac{L^2/T}{A}, \tag{2.37}$$

to compare advection and stochastic diffusion terms in the momentum equation, and will be referred as the *stochastic Peclet number*, in perfect similarity with the deterministic advection-diffusion problem (Quarteroni et al., 1994; Heinrich et al., 1977). The non dimensional stochastic transport operator is finally:

$$\hat{\mathbb{D}}_t\theta = \mathrm{d}_t\theta + \left\{\left[\hat{\boldsymbol{v}} - \frac{1}{\Upsilon}\boldsymbol{v}_s + \frac{1}{\Upsilon}\hat{\boldsymbol{\sigma}}_t^{\mathrm{T}}\left(\hat{\nabla}\cdot\hat{\boldsymbol{\sigma}}_t\right)\right]\mathrm{d}\hat{t} + \frac{\epsilon}{\Upsilon}\hat{\boldsymbol{\sigma}}_t\mathrm{d}\hat{\mathbf{B}}_t\right\}\cdot\hat{\nabla}\theta - \frac{1}{\Upsilon}\frac{1}{2}\hat{\nabla}\cdot\left(\hat{\boldsymbol{a}}\hat{\nabla}\theta\right)\mathrm{d}\hat{t}.$$

In the noise term, the scaling parameter Υ has been related to the ratio between the large scale and the small scales of the motion, that is $\epsilon = T_\sigma/T$, by using the definition of A as

$$\Upsilon = \frac{L^2/T}{A} = \frac{L^2/T}{L_\sigma^2/T_\sigma} = \epsilon^{-1}\frac{L_\sigma}{L}. \tag{2.38}$$

Notice that, introducing a uniform scaling parameter ε such that all the stochastic scales follow the same behaviour as

$$\varepsilon = \frac{L_\sigma}{L} = \frac{U_\sigma}{U} = \frac{T_\sigma}{T} \tag{2.39}$$

one finds the scaling

$$\hat{\mathbb{D}}_t\theta = \mathrm{d}_t\theta + \left\{\left[\hat{\boldsymbol{v}} - \frac{1}{\Upsilon}\boldsymbol{v}_s + \frac{1}{\Upsilon}\hat{\boldsymbol{\sigma}}_t^{\mathrm{T}}\left(\hat{\nabla}\cdot\hat{\boldsymbol{\sigma}}_t\right)\right]\mathrm{d}\hat{t} + \frac{1}{\Upsilon^{1/2}}\hat{\boldsymbol{\sigma}}_t\mathrm{d}\hat{\mathbf{B}}_t\right\}\cdot\hat{\nabla}\theta - \frac{1}{\Upsilon}\frac{1}{2}\hat{\nabla}\cdot\left(\hat{\boldsymbol{a}}\hat{\nabla}\theta\right)\mathrm{d}\hat{t}.$$

already described in the literature (Bauer et al., 2020b; Resseguier, Mémin, and Chapron, 2017b).

2.4 Conservation statements under location uncertainty

In this section, the expression of the volumetric rate given by Theorem 1 is used to establish balances and conservation statements to represent the conservation laws of fluid mechanics under LU, while the expression of the volumetric rate of change when stochastically forced provided by the Extended stochastic Reynolds transport of theorem 2 is also used to further generalise the results.

2.4.1 Conservation of mass

In this initial exploration of stochastic modelling, the conservation of mass is expressed within the stochastic setting. In the case of fluid mechanics, relativistic effects that can result in creation or destruction of mass can be neglected, so the conservation of mass inside a time-dependent volume V_t can be stated as

$$\mathrm{d}\int_{V_t}\rho\left(\boldsymbol{x},t\right)\,\mathrm{d}\boldsymbol{x}=0. \tag{2.40}$$

Considering the arbitrariness of the control volume V_t and using the localization theorem, mass conservation in differential form is

$$\mathrm{d}_t\rho+\nabla\cdot\left[\rho\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\rho\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\rho\right)\,\mathrm{d}t=0 \tag{2.41}$$

Alternatively, conservation of mass can be written in terms of the stochastic transport operator as:

$$\begin{gathered}\mathrm{d}_t\rho+\left[\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]\cdot\nabla\rho-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\rho\right)\,\mathrm{d}t+\rho\nabla\cdot\left[\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]=0\\ \mathbb{D}_t\rho+\rho\nabla\cdot\left[\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\boldsymbol{\sigma}_t^{\mathrm{T}}\left(\nabla\cdot\boldsymbol{\sigma}_t\right)\cdot\nabla\rho=0.\end{gathered} \tag{2.42}$$

Incompressible fluids

In the case of an incompressible fluid the mass conservation simplifies drastically. Imposing $\rho\left(\boldsymbol{x},t\right)$ =const, all the derivatives of ρ disappear, hence

$$\frac{\partial}{\partial x_k}\left(v^k-\frac{1}{2}\frac{\partial a_{ik}}{\partial x_i}\right)\,\mathrm{d}t+\frac{\partial}{\partial x_k}\int_\Omega \breve{\sigma}_{kj}\mathrm{d}\mathrm{B}_t^j\,\mathrm{d}\boldsymbol{y}=0. \tag{2.43}$$

Using the uniqueness of the semi-martingale decomposition Protter, 2004 one finds, setting the martingale contribution to zero:

$$\frac{\partial}{\partial x_k}\int_\Omega \breve{\sigma}_{kj}\mathrm{d}\mathrm{B}_t^j\,\mathrm{d}\boldsymbol{y}=0 \quad\Leftrightarrow\quad \frac{\partial}{\partial x_k}\breve{\sigma}_{kj}=0, \qquad \nabla\cdot\boldsymbol{\sigma}_t=0. \tag{2.44}$$

With this condition one can rearrange the integral term in the bounded variation term (alongside the symmetry of $\breve{\sigma}$) to get a modified incompressibility condition:

$$\frac{\partial}{\partial x_k}\left(v^k-\frac{1}{2}\frac{\partial}{\partial x_i}a_{ik}\right)=0, \qquad \nabla\cdot\left(\boldsymbol{v}-\boldsymbol{v}_s\right)=0. \tag{2.45}$$

Equation (2.44) provides a great insight to further modelling. When the fluid is assumed to be incompressible, the stochastic noise is itself incompressible, as it has a zero divergence.

2.4.2 Conservation of extensive tracers

The conservation of a scalar tracer is a generalization of the reasoning introduced in the previous section for conservation of mass. The starting balance is in the general case composed of a bounded variation and a martingale forcing,

$$\mathrm{d}\int_{V_t}\rho\theta\,\mathrm{d}\boldsymbol{x}=\int_{V_t}\left(\rho\Theta_t\,\mathrm{d}t+\rho\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t\right)\,\mathrm{d}\boldsymbol{x}, \tag{2.46}$$

so that the conservation of the scalar tracer is defined through Theorem 2. The balance in differential form reads thus:

$$\begin{aligned}\mathrm{d}_t\left(\rho\theta\right)+\nabla\cdot\left[\rho\theta\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\rho\theta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left[\boldsymbol{a}\nabla(\,\rho\theta\,)\right]\,\mathrm{d}t+\nabla\cdot\left(\boldsymbol{\sigma}_t\rho\boldsymbol{\Theta}_\sigma\right)\,\mathrm{d}t\\ =\rho\Theta_t\,\mathrm{d}t+\rho\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t.\end{aligned} \tag{2.47}$$

Applying conservation of mass, it is possible to write the previous conservation statement for $\rho\theta$ as a conservation statement for θ only, as

$$\begin{aligned} \mathrm{d}_t\theta + [(\boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}}\,(\nabla\cdot\boldsymbol{\sigma}_t))\,\mathrm{d}t + \boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t]\cdot\nabla\theta - \tfrac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}t + \nabla\cdot(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\,\mathrm{d}t \\ = \rho\Theta_t\,\mathrm{d}t + \rho\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t. \end{aligned} \tag{2.48}$$

or employing the stochastic transport operator, this reads

$$\mathbb{D}_t\theta + \mathrm{Tr}\,(\boldsymbol{\sigma}_t^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma)\,\mathrm{d}t = \Theta_t\,\mathrm{d}t + \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t. \tag{2.49}$$

It is worth remarking that in the previous equation the term $\mathrm{Tr}\,(\boldsymbol{\sigma}_t^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma)$ is a compact notation for the integral

$$\mathrm{Tr}\,(\boldsymbol{\sigma}_t^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma) = \int_\Omega \breve{\sigma}_{ik}\,(\mathbf{X}_t,\boldsymbol{y},t)\,\frac{\partial}{\partial x_i}\breve{\Theta}_\sigma^k\,(\mathbf{X}_t,\boldsymbol{y},t)\,\mathrm{d}\boldsymbol{y} \tag{2.50}$$

and it must be adapted case by case with the true definition of the martingale forcing $\boldsymbol{\Theta}_\sigma$.

2.4.3 Conservation of momentum

Conservation of momentum is stated as usual. Starting from Newton's second law for a volume V_t of fluid

$$\mathrm{d}\int_{V_t}\rho\boldsymbol{V}\,\mathrm{d}\boldsymbol{x} = \int_{V_t}\rho\boldsymbol{F}\,\mathrm{d}\boldsymbol{x}, \tag{2.51}$$

where $\rho\boldsymbol{F}$ represents the ensemble of the forces acting on the fluid volume V_t. The total velocity $\boldsymbol{V}$ is given in Equation (2.7), so by applying this definition it is clear that two different temporal scales exists. Considering Equation (2.51) integrated in time against a test function φ one has

$$\int_T\varphi\,(t)\left[\mathrm{d}\int_{V_t}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} + \mathrm{d}\int_{V_t}\rho\sigma\dot{\mathbf{B}}_t\,\mathrm{d}\boldsymbol{x}\right]\mathrm{d}t = \int_T\varphi\,(t)\int_{V_t}\rho\boldsymbol{F}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{2.52}$$

$\dot{\mathbf{B}}$ is expressed in the sense of distributions, that means

$$\int_T\varphi\,(t)\left[\mathrm{d}\int_{V_t}\rho\sigma\dot{\mathbf{B}}_t\,\mathrm{d}\boldsymbol{x}\right]\mathrm{d}t = -\int_T\varphi'\,(t)\left[\int_{V_t}\rho\sigma\mathrm{d}\mathbf{B}_t\,\mathrm{d}\boldsymbol{x}\right],$$

so that

$$\int_T \varphi\left(t\right)\left[\mathrm{d}\int_{V_t}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}\right]\,\mathrm{d}t-\int_T\varphi'\left(t\right)\left[\int_{V_t}\rho\sigma\mathrm{d}\mathbf{B}_t\,\mathrm{d}\boldsymbol{x}\right]=\int_T\varphi\left(t\right)\int_{V_t}\rho\boldsymbol{F}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{2.53}$$

In this equation, the first term is a semi-martingale (Kunita, 1997) term (loosely speaking, it contains both bounded variation terms and martingale terms) while the second is a Brownian acceleration expressed in weak sense, showing the existence of an extreme velocity scale. As justified rigorously by the pioneering work of Mikulevicius et al. (2004), the two sides of the previous equation must be consistent within their temporal scales. This means that $\boldsymbol{F}$, which remains thus far unspecified, must have different temporal scales

$$\int_T\varphi\left(t\right)\int_{V_t}\rho\boldsymbol{F}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t=\int_T\varphi\left(t\right)\left[\int_{V_t}\rho\left(\boldsymbol{F}_t\,\mathrm{d}t+\mathbf{F}_\sigma\mathrm{d}\mathbf{B}_t\right)\,\mathrm{d}\boldsymbol{x}\right]\,\mathrm{d}t-\int_T\varphi'\left(t\right)\left[\int_{V_t}\rho\ddot{\mathbf{F}}\mathrm{d}\mathbf{B}_t\,\mathrm{d}\boldsymbol{x}\right].$$

The last term $\ddot{\mathbf{F}}\mathrm{d}\mathbf{B}_t$, associated with the Brownian acceleration, has to be considered as a balancing term for the very-fast scale introduced in (2.51) by the temporal derivative of the velocity $\boldsymbol{V}=\boldsymbol{v}+\sigma\dot{\mathbf{B}}_t$, that already contains a white noise term. For its interpretation as a Brownian acceleration, double dot notation has been chosen (However, it does not correspond to a double differentiation of $\boldsymbol{F}$, it is only a notation). Assuming that the balance between the *very fast* scales associated to the Brownian acceleration and the corresponding *very fast* forces holds, the final momentum balance reads

$$\mathrm{d}\int_{V_t}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=\int_{V_t}\rho\left(\boldsymbol{F}_t\,\mathrm{d}t+\rho\mathbf{F}_\sigma\mathrm{d}\mathbf{B}_t\right)\,\mathrm{d}\boldsymbol{x}, \tag{2.54}$$

where the first term is a bounded variation forcing term, the second can be defined in general as a martingale forcing term:

$$\mathrm{F}_\sigma\mathrm{d}\mathrm{B}_t^i=\int_\Omega\breve{F}_\sigma^{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{d}\mathrm{B}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y},\quad i=1,2,3. \tag{2.55}$$

Applying the extended Reynolds' transport theorem one finds

$$\begin{aligned}\int_{V_t}\rho\Big[\mathrm{d}_t\left(\rho\boldsymbol{v}\right)+\nabla\cdot\left[\rho\boldsymbol{v}\otimes\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\rho\boldsymbol{v}\otimes\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left[\boldsymbol{a}\nabla\left(\rho\boldsymbol{v}\right)\right]\,\mathrm{d}t\\+\nabla\cdot\left(\rho\boldsymbol{\sigma}_t\mathbf{F}_\sigma\right)\,\mathrm{d}t\right]\,\mathrm{d}\boldsymbol{x}=\int_{V_t}\rho\left(\boldsymbol{F}_t\,\mathrm{d}t+\mathbf{F}_\sigma\mathrm{d}\mathbf{B}_t\right)\,\mathrm{d}\boldsymbol{x},\end{aligned} \tag{2.56}$$

where $\otimes$ denotes the outer product. Moreover, the previous balance can be written using the distributivity as

$$\int_{V_t} \rho \Big[\mathrm{d}_t\theta + [\boldsymbol{v}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \,\mathrm{d}\mathbf{B}_t] \cdot \nabla \boldsymbol{v} - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla \boldsymbol{v}) \,\mathrm{d}t \\ + \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \mathbf{F}_\sigma\right) \mathrm{d}t - \boldsymbol{F}_t \,\mathrm{d}t - \mathbf{F}_\sigma \mathrm{d}\mathbf{B}_t] \,\mathrm{d}\boldsymbol{x} = 0$$

with $\boldsymbol{v}^\star = \boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma}_t\right)$. Using the definition of the stochastic transport operator one can rearrange the terms as

$$\int_{V_t} \rho \left[\mathbb{D}_t \boldsymbol{v} + \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \mathbf{F}_\sigma\right) \mathrm{d}t - \boldsymbol{F}_t \,\mathrm{d}t - \mathbf{F}_\sigma \mathrm{d}\mathbf{B}_t\right] \mathrm{d}\boldsymbol{x} = 0.$$

Finally, the differential form of momentum balance reads

$$\rho \mathbb{D}_t \boldsymbol{v} + \rho \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \mathbf{F}_\sigma\right) \mathrm{d}t = \rho \left(\boldsymbol{F}_t \,\mathrm{d}t + \mathbf{F}_\sigma \mathrm{d}\mathbf{B}_t\right).$$

Note that in many situations the correlation term $\rho \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \mathbf{F}_\sigma\right)$ is neglected ant the form

$$\rho \mathbb{D}_t \boldsymbol{v} = \rho \left(\boldsymbol{F}_t \,\mathrm{d}t + \mathbf{F}_\sigma \mathrm{d}\mathbf{B}_t\right).$$

is obtained.

2.4.4 Conservation of energy

The conservation of energy is a generalization of the reasoning introduced in the previous section for conservation of extensive tracer. Defining with $\mathrm{d}W_t$ and $\boldsymbol{W}_\sigma \cdot \mathrm{d}\mathbf{B}_t$ the bounded variation and the martingale part of the work exerced on the system, the starting balance is, in the general case of a finite variation and a martingale forcing,

$$\mathrm{d} \int_{V_t} \rho E \,\mathrm{d}\boldsymbol{x} = \int_{V_t} \rho \left(\mathrm{d}W_t + \boldsymbol{W}_\sigma \cdot \mathrm{d}\mathbf{B}_t\right) \mathrm{d}\boldsymbol{x}, \tag{2.57}$$

so that the conservation of the scalar tracer is defined through Theorem 2, and the balance in differential reads:

$$\mathrm{d}_t \left(\rho E\right) + \nabla \cdot \left[\rho E \left(\boldsymbol{v} - \boldsymbol{v}_s\right) \mathrm{d}t + \rho E \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t\right] - \frac{1}{2}\nabla \cdot \left(\boldsymbol{a}\nabla \rho E\right) \mathrm{d}t + \nabla \cdot \left(\boldsymbol{\sigma}_t \rho \boldsymbol{W}_\sigma\right) \mathrm{d}t \\ = \rho \mathrm{d}W_t + \rho \boldsymbol{W}_\sigma \cdot \mathrm{d}\mathbf{B}_t. \tag{2.58}$$

Applying conservation of mass, in the same way it was done for momentum, it is possible to write the previous conservation statement for ρE as a conservation statement for E only, as

$$\begin{aligned} \mathrm{d}_t E + [(\boldsymbol{v} - \boldsymbol{v}_s + \boldsymbol{\sigma}_t^{\mathrm{T}}(\nabla \cdot \boldsymbol{\sigma}_t))\,\mathrm{d}t + \boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t] \cdot \nabla E - \tfrac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla E)\,\mathrm{d}t + \nabla \cdot (\boldsymbol{\sigma}_t \boldsymbol{W}_\sigma)\,\mathrm{d}t \\ = \rho \mathrm{d}W_t + \rho \boldsymbol{W}_\sigma \cdot \mathrm{d}\mathbf{B}_t, \end{aligned} \tag{2.59}$$

which can be compactly expressed using the stochastic transport operator as

$$\mathbb{D}_t E + \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \boldsymbol{W}_\sigma\right)\mathrm{d}t = \rho\left(\mathrm{d}W_t + \boldsymbol{W}_\sigma \cdot \mathrm{d}\mathbf{B}_t\right). \tag{2.60}$$

2.5 Stochastic Navier-Stokes equations (SNS)

Navier-Stokes equations represent the conservation of mass and of momentum for a Newtonian fluid. The stress tensor for a Newtonian fluid is thus

$$\boldsymbol{T}_{ij} = -\left(p + \frac{2}{3}\mu\frac{\partial v^k}{\partial x_k}\right)\delta_{ij} + \mu\left(\frac{\partial v^i}{\partial x_j} + \frac{\partial v^j}{\partial x_i}\right), \tag{2.61}$$

where $p = -\frac{1}{3}\boldsymbol{T}_{ii}$ is the dynamical pressure, defining the isotropic component of the stress tensor, while the remaining terms (usually grouped into a τ_{ij} term) correspond to the deviatoric stress tensor, that is entirely determined by the motion of the fluid itself, as it describes entirely the difference with respect to the equilibrium state. Cauchy's equation of motions relates the variation of momentum to the divergence of the stress tensor, so the impulse is given by

$$\frac{\partial}{\partial x_j}\boldsymbol{T}_{ij} = -\frac{\partial p}{\partial x_i} + \mu\frac{\partial^2 v^i}{\partial x_j \partial x_j} + \frac{\mu}{3}\frac{\partial}{\partial x_i}\frac{\partial v^k}{\partial x_k}. \tag{2.62}$$

In the stochastic setting, the velocity $\boldsymbol{v}$ is no more deterministic. Hence the stress tensor $\boldsymbol{T}$ must be adapted to the new formulation. The divergence of the stress tensor, *i.e.* the impulse of the forces acting on the fluid (without considering body forces such as gravity) is :

$$\mathrm{d}\mathcal{J}_i = -\frac{\partial}{\partial x_i}\left(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma\right) + \mu\frac{\partial^2}{\partial x_j \partial x_j}\left(v^i \mathrm{d}t + \sigma \mathrm{dB}_t^i\right) + \frac{\mu}{3}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_k}\left(v^k \mathrm{d}t + \sigma \mathrm{dB}_t^k\right) \tag{2.63}$$

$$= \left[-\frac{\partial p}{\partial x_i} + \mu\frac{\partial^2 v^i}{\partial x_j \partial x_j} + \frac{\mu}{3}\frac{\partial}{\partial x_i}\frac{\partial v^k}{\partial x_k}\right]\mathrm{d}t - \frac{\partial}{\partial x_i}\mathrm{d}p_t^\sigma + \mu\frac{\partial^2}{\partial x_j \partial x_j}\sigma \mathrm{dB}_t^i \tag{2.64}$$

$$+\frac{\mu}{3}\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_k}\sigma \mathrm{dB}_t^k\right),$$

where (2.7) was employed and where the unique decomposition of semi-martingale property was used in the last equality. In vector notation one can write

$$\mathrm{d}\boldsymbol{\mathcal{J}} = \left(-\nabla p + \rho \boldsymbol{F}_t\right)\,\mathrm{d}t + \left(-\mathrm{d}p_t^\sigma + \rho \mathbf{F}_\sigma \mathrm{d}\mathbf{B}_t\right), \tag{2.65}$$

where the following notation was introduced to simplify the exposition:

$$\begin{aligned}\boldsymbol{F}_t &= \nabla\left(\frac{\mu}{3}\nabla\cdot\boldsymbol{v}\right) + \mu\Delta\boldsymbol{v},\\ \mathbf{F}_\sigma \mathrm{d}\mathbf{B}_t &= \nabla\left(\frac{\mu}{3}\nabla\cdot\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right) + \mu\Delta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t.\end{aligned}$$

Stochastic Navier-Stokes equations can finally be formulated as the combination of the conservation of momentum

$$\begin{aligned}\mathbb{D}_t\boldsymbol{v} - \nabla\cdot\left[\left(\nabla\boldsymbol{\vartheta}\right)\boldsymbol{\sigma}_t\right]\mathrm{d}t + \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}}\nabla\mathbf{F}_\sigma\right)\,\mathrm{d}t =& \frac{1}{\rho}\left[\nabla\left(-p + \frac{\mu}{3}\nabla\cdot\boldsymbol{v}\right) + \mu\Delta\boldsymbol{v}\right]\mathrm{d}t\\ &+ \frac{1}{\rho}\nabla\left(-\mathrm{d}p_t^\sigma + \frac{\mu}{3}\nabla\cdot\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right) + \mu\Delta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t,\end{aligned}$$

and conservation of mass

$$\mathrm{d}_t\rho + \nabla\cdot\left[\rho\left(\boldsymbol{v} - \boldsymbol{v}_s\right)\,\mathrm{d}t + \rho\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right] - \frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\rho\right)\,\mathrm{d}t = 0,$$

usually complemented with an equation of state. The meaning of the term $\nabla\cdot\left[\left(\nabla\boldsymbol{\vartheta}\right)\boldsymbol{\sigma}_t\right]\mathrm{d}t$ is detailed in the next Section.

Stochastic Navier-Stokes equations

Momentum equation: (2.66)

$$\mathbb{D}_t \boldsymbol{v} - \nabla \cdot \left[(\nabla \boldsymbol{\vartheta}) \, \boldsymbol{\sigma}_t \right] \mathrm{d}t + \mathrm{Tr} \left(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \mathbf{F}_\sigma \right) \, \mathrm{d}t = \left[\frac{1}{\rho} \nabla \left(-p + \frac{\mu}{3} \nabla \cdot \boldsymbol{v} \right) + \nu \Delta \boldsymbol{v} \right] \mathrm{d}t + \frac{1}{\rho} \nabla \left(-\mathrm{d}p_t^\sigma + \frac{\mu}{3} \nabla \cdot \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \right) + \nu \Delta \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t,$$

Mass transport: (2.67)

$$\mathrm{d}_t \rho + \nabla \cdot \left[\rho \left(\boldsymbol{v} - \boldsymbol{v}_s \right) \, \mathrm{d}t + \rho \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \right] - \frac{1}{2} \nabla \cdot \left(\boldsymbol{a} \nabla \rho \right) \, \mathrm{d}t = 0,$$

Equation of state: (2.68)

$$\rho = \rho \left(T, p \right).$$

Stochastic pressure

Equation (2.65) gives an important formulation for the impulse acting on a flow in the stochastic setting. However, the term $\mathrm{d}p_t^\sigma$ does not seem to appear naturally but rather through a modelling choice. This term is the *stochastic pressure*, defined as

$$\mathrm{d}p_t^\sigma \left(\boldsymbol{x} \right) = \int_\Omega \vartheta_i \left(\boldsymbol{x}, \boldsymbol{y}, t \right) \mathrm{dB}_t^i \left(\boldsymbol{y} \right) \mathrm{d}\boldsymbol{y}, \tag{2.69}$$

and is a zero-mean turbulent pressure related to the small scale velocity component. The motivating argument behind the introduction of this term is that equations should be consistent in their time-scale separation. Consider the stochastic Euler equations, that is the inviscid version of the previously introduced Navier-Stokes equations, forced by a generic pressure dP_t. Applying the divergence operator to the momentum equation one finds

$$\mathbb{D}_t \left(\frac{\partial \boldsymbol{v}}{\partial x_k} \right) + \frac{\partial}{\partial x_k} \left[\boldsymbol{v}^\star \, \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \right]_j \left(\frac{\partial \boldsymbol{v}}{\partial x_j} \right) - \frac{1}{2} \frac{\partial}{\partial x_i} \left(\frac{\partial a_{ij}}{\partial x_k} \frac{\partial \boldsymbol{v}}{\partial x_j} \right) = \Delta \mathrm{dP}_t. \tag{2.70}$$

As the right and left hand sides must match in both slow scale terms and fast scale terms, the right hand side must have both a bounded variation and a purely stochastic term.

Hence, the total pressure $\mathrm{d}\mathrm{P}_t$ is split into two components

$$\mathrm{d}\mathrm{P}_t = p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}. \tag{2.71}$$

The second term, the martingale component of the pressure, has been split in Equation (2.65) and it is responsible for the term $\nabla \cdot [(\nabla \boldsymbol{\vartheta})\,\boldsymbol{\sigma}_t]\,\mathrm{d}t$ in Equation (2.66), as detailed in Section 2.9.5. In practice, both terms $\nabla \cdot [(\nabla \boldsymbol{\vartheta})\,\boldsymbol{\sigma}_t]\,\mathrm{d}t$ and $\mathrm{Tr}\,(\boldsymbol{\sigma}^{\mathrm{T}} \nabla \mathbf{F}_{\sigma})$ are often neglected, recovering the models described in Resseguier, Mémin, Heitz, et al. (2017).

2.6 Stochastic Boussinesq equations

Density variations are often induced by temperature or chemicals (as salinity), but these effects can be often regarded as small. To deal with the results of these changes without using the compressible Navier-Stokes equations, the so called *Boussinesq approximation* is employed. This approximation assumes that the density field involves just small changes around a reference ρ_0, so that the density field can be written as

$$\begin{aligned} \rho\,(\boldsymbol{x}, t) &= \rho_0 + \delta\rho\,(\boldsymbol{x}, t) \\ &= \rho_0\,(1 + \epsilon\,\delta\hat{\rho}\,(\boldsymbol{x}, t))\,, \end{aligned} \tag{2.72}$$

where $\delta\hat{\rho}\,(\boldsymbol{x}, t)$ is a non-dimensional variable and $\epsilon \ll 1$ is a constant small parameter. In the deterministic case, starting from the continuity equation

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \nabla \cdot \boldsymbol{v} = 0 \tag{2.73}$$

and applying Equation (2.72) one has

$$\nabla \cdot \boldsymbol{v} + \epsilon \left[\frac{\mathrm{D}\delta\hat{\rho}}{\mathrm{D}t} + \delta\hat{\rho} \nabla \cdot \boldsymbol{v} \right] = 0. \tag{2.74}$$

It is clear that the first order term in this equation involves just the isochoric constraint. If ϵ is small, the incompressibility condition is sufficient. This of course does not imply that the fluid is itself incompressible, but rather that mathematically at the limit the flow can be considered as isochoric. In the LU model the modification of the conservation of

mass follows the same path and the result is

$$\nabla\cdot[(\boldsymbol{v}-\boldsymbol{v}_s)\ \mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t]+\epsilon\left[\mathrm{d}_t\delta\hat{\rho}+\nabla\cdot[\delta\hat{\rho}\,(\boldsymbol{v}-\boldsymbol{v}_s)\ \mathrm{d}t+\delta\hat{\rho}\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t]-\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\,\delta\hat{\rho})\ \mathrm{d}t\right]=0. \tag{2.75}$$

At first order, the divergence-free condition

$$\nabla\cdot[(\boldsymbol{v}-\boldsymbol{v}_s)\ \mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t]=0 \tag{2.76}$$

is found. This implies, separating the scales of motion, that the noise term is incompressible by itself, with

$$\nabla\cdot\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t=0 \tag{2.77}$$

implying that $\nabla\cdot\boldsymbol{\sigma}_t=0$, so that the modified advection reduces to $\boldsymbol{v}^\star=\boldsymbol{v}-\boldsymbol{v}_s$. This result is analogous to what obtained for incompressible fluids in Section 2.4.1. The advective form can be derived directly from the stochastic Reynolds transport theorem where the term $\boldsymbol{\sigma}^{\mathrm{T}}\,(\nabla\cdot\boldsymbol{\sigma}_t)$ cancels, meaning that the two forms

$$\mathrm{d}_t\theta+\nabla\cdot[\theta\,(\boldsymbol{v}-\boldsymbol{v}_s)\ \mathrm{d}t+\theta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t]-\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\ \mathrm{d}t+\nabla\cdot(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\ \mathrm{d}t=\Theta_t\,\mathrm{d}t+\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t,$$
$$\mathrm{d}_t\theta+[(\boldsymbol{v}-\boldsymbol{v}_s)\ \mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t]\cdot\nabla\theta-\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\ \mathrm{d}t+\nabla\cdot(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\ \mathrm{d}t=\Theta_t\,\mathrm{d}t+\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t,$$

are interchangeable. Moreover, the continuity equation states that at lowest order there is no dynamics for the density. Conservation of momentum can be derived starting from an integral balance between the variation of momentum, the gravitational field acting on the fluid and the divergence of the stress tensor for a Newtonian fluid, neglecting terms involving molecular diffusion (as a standard practice in large scale dynamics) and noise compressibility, resulting in:

$$\mathrm{d}\int_{V_t}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=-\int_{V_t}[(\nabla p+\rho g\mathbf{e}_z)\ \mathrm{d}t+\nabla\mathrm{d}p_t^\sigma]\ \mathrm{d}\boldsymbol{x}. \tag{2.78}$$

Inserting (2.72) one finds

$$\mathrm{d}\int_{V_t}(\rho_0+\delta\rho)\,\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=-\int_{V_t}[(\nabla p+(\rho_0+\delta\rho)\,g\mathbf{e}_z)\ \mathrm{d}t+\nabla\mathrm{d}p_t^\sigma]\ \mathrm{d}\boldsymbol{x}, \tag{2.79}$$

that can be rearranged conveniently as

$$\mathrm{d}\int_{V_t}\rho_0\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}+\mathrm{d}\int_{V_t}\delta\rho\,\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=\int_{V_t}-\nabla\left[p\,\mathrm{d}t+\mathrm{d}p_t^\sigma\right]\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}-\int_{V_t}\rho_0 g\mathbf{e}_z\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}-\int_{V_t}\delta\rho g\mathbf{e}_z\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}$$
$$\mathrm{d}\int_{V_t}\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}+\mathrm{d}\int_{V_t}\frac{\delta\rho}{\rho_0}\,\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=\int_{V_t}-\frac{1}{\rho_0}\nabla\left[p\,\mathrm{d}t+\mathrm{d}p_t^\sigma\right]\,\mathrm{d}\boldsymbol{x}-\int_{V_t}g\mathbf{e}_z\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}-\int_{V_t}\frac{\delta\rho}{\rho_0}g\mathbf{e}_z\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}.$$

To understand the leading orders of the momentum equation one must employ an asymptotic analysis, so the non-dimensional variables

$$\boldsymbol{x}=L\hat{\boldsymbol{x}},\qquad \boldsymbol{v}=U\hat{\boldsymbol{v}},\qquad t=\frac{L}{U}\hat{t},\qquad p=P\hat{p},\qquad g=\frac{P}{\rho_0 L}\hat{g}. \tag{2.80}$$

are introduced, with the definition $\frac{\delta\rho}{\rho_0}=\epsilon\hat{\rho}_1$, so that the non dimensional integral momentum balance reads

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}+\epsilon\mathrm{d}\int_{\hat{V}_t}\hat{\rho}_1\,\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}=-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\hat{\nabla}\left[\frac{PL}{U}\hat{p}\,\mathrm{d}\hat{t}+P^\sigma\mathrm{d}\hat{p}_t^\sigma\right]+\frac{PL}{U}\hat{g}\mathbf{e}_3\,\mathrm{d}\hat{t}+\epsilon\frac{PL}{U}\hat{\rho}_1\hat{g}\mathbf{e}_3\,\mathrm{d}\hat{t}\,\mathrm{d}\hat{\boldsymbol{x}}.$$

The scaling P^σ for the martingale part of the pressure was also introduced. Expanding each variable [1] as an asymptotic with ϵ taken as a smallness parameter, that is

$$\phi=\phi_0+\epsilon\phi_1+\epsilon^2\phi_2+\epsilon^3\phi_3+\dots \tag{2.81}$$

with ϕ representing either the pressure, velocity or variation of density, one has

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}_0\,\mathrm{d}\hat{\boldsymbol{x}}+\epsilon\mathrm{d}\int_{\hat{V}_t}\hat{\rho}_1\,\hat{\boldsymbol{v}}_0\,\mathrm{d}\hat{\boldsymbol{x}}=-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\left[\frac{PL}{U}\left[\hat{\nabla}\left(\hat{p}_0+\epsilon\hat{p}_1\right)+\hat{g}\right]\,\mathrm{d}\hat{t}\right.$$
$$\left.+\,\hat{\nabla}P_0^\sigma\mathrm{d}\hat{p}_t^\sigma+\epsilon\frac{PL}{U}\hat{\rho}_1\hat{g}\mathbf{e}_3\,\mathrm{d}\hat{t}\right]\,\mathrm{d}\hat{\boldsymbol{x}}$$

defining $P=\rho^\star U^2$ with $\rho^\star$ such that $\rho^\star=\dfrac{\rho_0}{\epsilon}$ provides

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}+\epsilon\mathrm{d}\int_{\hat{V}_t}\hat{\rho}_1\,\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}=\frac{1}{\epsilon}\int_{\hat{V}_t}\left[\hat{\nabla}\hat{p}_0+\hat{g}\mathbf{e}_3\right]\,\mathrm{d}\hat{\boldsymbol{x}}-\int_{\hat{V}_t}\left[\hat{\nabla}\hat{p}_1+\hat{\rho}_1\hat{g}\mathbf{e}_3\right]\,\mathrm{d}\hat{t}\,\mathrm{d}\hat{\boldsymbol{x}}$$
$$-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\hat{\nabla}P^\sigma\mathrm{d}\hat{p}_t^\sigma\,\mathrm{d}\hat{\boldsymbol{x}}.$$

1. Theoretically, one should expand each variable, $\hat{\boldsymbol{v}}$, $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{\sigma}}_t$, $\hat{p}$, $\delta\hat{\rho}$. In practice, the main result of this section can be achieved by expanding the pressure to first order as $p=p_0+\epsilon p_1+\mathcal{O}\left(\epsilon^2\right)$ while all the remaining terms can be taken as zero-th order $\phi=\phi_0+\mathcal{O}\left(\epsilon\right)$

This choice, that might seem *ad hoc* at first glimpse, is stating that the dynamics is driven by density variations rather than the mean state, thus the reference density is defined as the mean state density ρ_0 amplified by the ratio between the mean state and the variation, $\frac{\rho_0}{\delta\rho} = \frac{1}{\epsilon}$. In this equation, the leading order is $\mathcal{O}(1/\epsilon)$ and is described by a hydrostatic distribution of the pressure

$$\int_{\hat{V}_t} \nabla \hat{p}_0 \, \mathrm{d}\hat{\boldsymbol{x}} = -\int_{\hat{V}_t} \hat{g}\boldsymbol{e}_3 \, \mathrm{d}\hat{\boldsymbol{x}}. \tag{2.82}$$

Once dimensional variables are replaced to non-dimensional variables into equation (2.82), the hydrostatic balance reads

$$\nabla p_0 = -\rho_0 g \boldsymbol{e}_3 \tag{2.83}$$

and so pressure has a zero-th order component that is only depending on the vertical coordinate and on the mean state density:

$$p_0(z) = -\rho_0 g z. \tag{2.84}$$

The separation of the density into two components, the background constant density and the deviation from it, corresponds thus to a division of the pressure between hydrostatic and the fluctuation from this reference,

$$\rho(\boldsymbol{x}, t) = \rho_0 + \delta\rho(\boldsymbol{x}, t), \qquad p(\boldsymbol{x}, t) = -\rho_0 g z + p'(\boldsymbol{x}, t) \tag{2.85}$$

with this latter fluctuation that can be identified with the first order component of the pressure

$$p'(\boldsymbol{x}, t) = p(\boldsymbol{x}, t) + \rho_0 g z. \tag{2.86}$$

The following order, $\mathcal{O}(\epsilon)$, reads

$$\mathrm{d}\int_{\hat{V}_t} \hat{\boldsymbol{v}} \, \mathrm{d}\hat{\boldsymbol{x}} = -\int_{\hat{V}_t} \left[\hat{\nabla}\hat{p}_1 + \hat{\rho}_1 \hat{g}\mathbf{e}_3\right] \mathrm{d}\hat{t} \, \mathrm{d}\hat{\boldsymbol{x}} - \frac{1}{\rho_0 U L}\int_{\hat{V}_t} \hat{\nabla} P^\sigma \mathrm{d}\hat{p}_t^\sigma, \tag{2.87}$$

so the stochastic Reynolds transport theorem can be applied to this equation, providing

$$\begin{aligned} \mathrm{d}_t \boldsymbol{v} + \nabla \cdot \left[\boldsymbol{v} \otimes (\boldsymbol{v} - \boldsymbol{v}_s) \, \mathrm{d}t + \boldsymbol{v} \otimes \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t\right] - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla \boldsymbol{v}) \, \mathrm{d}t \\ -\nabla \cdot \left[(\nabla \boldsymbol{\vartheta}) \, \boldsymbol{\sigma}_t\right] \mathrm{d}t = -\frac{1}{\rho_0}\nabla \left[p' \, \mathrm{d}t + \mathrm{d}p_t^\sigma\right] - \frac{\delta\rho}{\rho_0} g \boldsymbol{e}_3 \, \mathrm{d}t. \end{aligned} \tag{2.88}$$

The last term is the *buoyancy* and depends on the density fluctuations, as indeed

$$\boldsymbol{b}\left(\boldsymbol{x},t\right)=-g\frac{\delta\rho\left(\boldsymbol{x},t\right)}{\rho_0}\boldsymbol{e}_3, \tag{2.89}$$

representing the upward (or downward) force associated with the density anomaly $\delta\rho$. In terms of buoyancy the Stochastic Boussinesq equations can be written as

$$\mathbb{D}_t\boldsymbol{v}-\nabla\cdot\left[\left(\nabla\boldsymbol{\vartheta}\right)\boldsymbol{\sigma}_t\right]\mathrm{d}t=-\frac{1}{\rho_0}\nabla\left[p'\,\mathrm{d}t+\mathrm{d}p_t^{\sigma}\right]-\boldsymbol{b}\,\mathrm{d}t, \tag{2.90}$$

where the stochastic transport operator has no $\boldsymbol{\sigma}_t^{\mathrm{T}}\left(\nabla\cdot\boldsymbol{\sigma}_t\right)$ term due to incompressibility of the noise (at first order). Considering now the density variation $\delta\,\rho$, order $\mathcal{O}\left(\epsilon\right)$ of the conservation of mass statement provides

$$\mathrm{d}_t\delta\hat{\rho}+\nabla\cdot\left[\delta\hat{\rho}\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\delta\hat{\rho}\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\,\delta\hat{\rho}\right)\,\mathrm{d}t=0, \tag{2.91}$$

that after multiplication by $\frac{g}{\rho_0}$ provides an equation for transport of buoyancy:

$$\mathbb{D}_t\,b=0. \tag{2.92}$$

A usual form of the buoyancy equation is retrieved straightforwardly when considering the effect of the mean stratification. In this case the density fluctuation $\delta\rho$ is split into a reference stratification $\tilde{\rho}\left(z\right)$ and its deviation from it, $\rho'\left(x,y,z,t\right)$, in such a way that

$$\rho\left(\boldsymbol{x},t\right)=\rho_0+\tilde{\rho}\left(z\right)+\rho'\left(x,y,z,t\right), \tag{2.93}$$

so that the buoyancy effects are now identified with ρ', meaning that buoyancy is defined as

$$\boldsymbol{b}\left(\boldsymbol{x},t\right)=-g\frac{\rho'\left(\boldsymbol{x},t\right)}{\rho_0}\boldsymbol{e}_3. \tag{2.94}$$

The resulting equation is

$$\begin{aligned}\mathrm{d}_t\rho'+\nabla\cdot\left[\rho'\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\rho'\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\,\rho'\right)\,\mathrm{d}t\\ +\nabla\cdot\left[\tilde{\rho}\left(\boldsymbol{v}-\boldsymbol{v}_s\right)\,\mathrm{d}t+\tilde{\rho}\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]-\frac{1}{2}\nabla\cdot\left(\boldsymbol{a}\nabla\,\tilde{\rho}\right)\,\mathrm{d}t=0,\end{aligned}$$

that can be written as

$$\mathbb{D}_t\left(b+N^2z\right)=0, \tag{2.95}$$

where the buoyancy frequency (or Brunt-Väisälä frequency) $N^2 = -\frac{g}{\rho_0}\frac{\mathrm{d}\tilde{\rho}}{\mathrm{d}z}$ was introduced.

Stochastic Boussinesq equations

Momentum equation: (2.96)

$$\mathbb{D}_t \boldsymbol{v} - \nabla \cdot [(\nabla \boldsymbol{\vartheta})\, \boldsymbol{\sigma}_t]\, \mathrm{d}t = -\frac{1}{\rho_0}\nabla \left[p'\, \mathrm{d}t + \mathrm{d}p_t^{\sigma}\right] - \boldsymbol{b}\, \mathrm{d}t,$$

Buoyancy transport: (2.97)

$$\mathbb{D}_t \left(b + N^2 z\right) = 0,$$

Incompressibility: (2.98)

$$\nabla \cdot [\boldsymbol{v} - \boldsymbol{v}_s] = 0, \qquad \nabla \cdot \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = 0.$$

2.7 Stochastic geophysical flows

Dealing with geophysical problems adds complexity even in the standard deterministic setting. These complexities arise from considering a spherical coordinate system, the geographical coordinate system composed of *Longitude* λ (the eastward rotation with respect to the conventional zero-meridian), *Latitude* θ (the northward rotation from the conventional equatorial plane) and the distance from the centre r, and considering the fictitious forces that appear when using this non-inertial system. In the stochastic setting the major issue presented is the definition of a temporal variation of a vector rotating around a given axis, but the increment of the vector can still be represented in terms of Frenet–Serret formulas:

$$\mathrm{d}\,\boldsymbol{C} = |C|\ |\Omega\, \mathrm{d}t|\, (\sin\theta)\, \mathbf{m} \tag{2.99}$$

where $\mathbf{m}$ is direction of the circular motion of the point C connected with the centre by the vector $\boldsymbol{C}$. This formula can be written in vector notation as $\mathrm{d}\boldsymbol{C} = \boldsymbol{\Omega}\, \mathrm{d}t \times \boldsymbol{C}$. In spherical coordinates the unit vectors are $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, corresponding to the *eastward*, *northward* and the *radially outward* vectors, as illustrated in Figure 2.1. The directions of these vectors change with position, and thus their derivatives are not zero. In the classical deterministic setting this rotation induces some modifications in the equation of motions, usually referred as *metric terms*. To compute these terms one proceeds computing $\boldsymbol{\Omega}_f$,

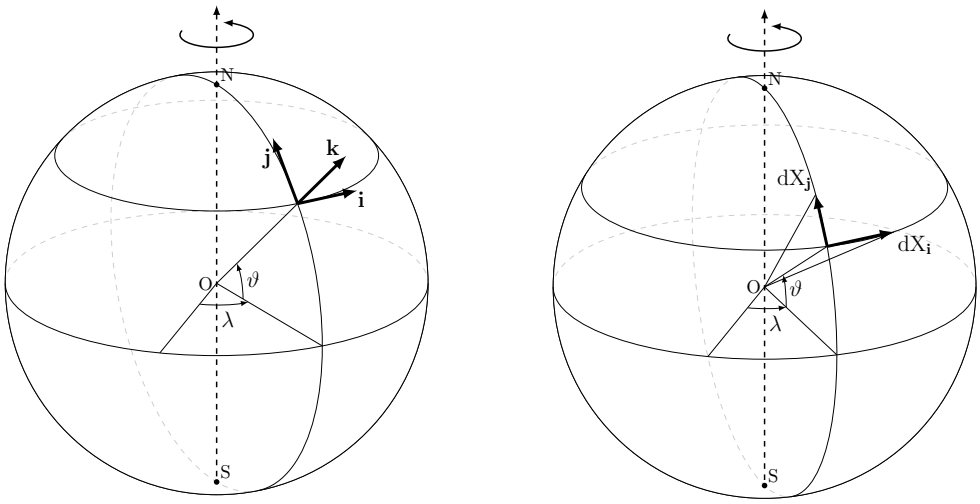


Figure 2.1 – Representation of the geographical coordinate system and associated tangent coordinate system.

the rotation with respect to the earth induced by the motion of the flow and use this rotation to define the metric terms, as done in Vallis, 2017. The usual procedure is that of using the flow $\boldsymbol{v} = (u, v, w)$ to compute the displacements induced in a time interval δt. In the LU framework however, the displacements are already defined explicitly as (2.1), thus considering the displacements along $\mathbf{i}$ and $\mathbf{j}$ (the only two displacements that cause a change in the frame) one has

$$r\,\mathrm{d}\theta = \mathrm{dX}_{\mathbf{j}} \tag{2.100}$$

$$\mathrm{d}\lambda\, r\cos\theta = \mathrm{dX}_{\mathbf{i}} \tag{2.101}$$

where the first, inducing a variation in latitude $\mathrm{d}\theta$, is a negative rotation around the $\mathbf{i}$ axis, while the second is a rotation around the axis of the earth, as it induces a variation in in longitude $\mathrm{d}\lambda$. This rotation around earth axis can be written as

$$\boldsymbol{\Omega}\,\mathrm{d}t = \Omega\,\mathrm{d}t\,(\mathbf{j}\cos\theta + \mathbf{k}\sin\theta) \tag{2.102}$$

since the rotation vector $\boldsymbol{\Omega}$ completely lies in the $(\mathbf{i}, \mathbf{k})$ plane. The infinitesimal displacement $\Omega\,\mathrm{d}t$ can be regarded as the rotation induced by the zonal motion, hence $\Omega\,\mathrm{d}t = \mathrm{dX}_{\mathbf{i}}/r\cos\theta$, i.e.

$$\boldsymbol{\Omega}_f\,\mathrm{d}t = -\frac{\mathrm{dX}_{\mathbf{j}}}{r}\mathbf{i} + \frac{\mathrm{dX}_{\mathbf{i}}}{r}\mathbf{j} + \frac{\mathrm{dX}_{\mathbf{i}}\tan\theta}{r}\mathbf{k} \tag{2.103}$$

where for the rotation around $\mathbf{i}$ the rotation induced by the meridional motion $\mathrm{d}\theta = \mathrm{dX}_\mathbf{j}/r$ was used. The variations of the vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ can then be defined as

$$\mathrm{d}\mathbf{i} = \boldsymbol{\Omega}_f \mathrm{d}t \times \mathbf{i} = \frac{\mathrm{dX}_\mathbf{i} \tan\theta}{r}\mathbf{j} + \frac{\mathrm{dX}_\mathbf{i}}{r}\mathbf{k}, \tag{2.104}$$

$$\mathrm{d}\mathbf{j} = \boldsymbol{\Omega}_f \mathrm{d}t \times \mathbf{j} = -\frac{\mathrm{dX}_\mathbf{i} \tan\theta}{r}\mathbf{i} - \frac{\mathrm{dX}_\mathbf{j}}{r}\mathbf{k}, \tag{2.105}$$

$$\mathrm{d}\mathbf{k} = \boldsymbol{\Omega}_f \mathrm{d}t \times \mathbf{k} = \frac{\mathrm{dX}_\mathbf{i}}{r}\mathbf{i} + \frac{\mathrm{dX}_\mathbf{j}}{r}\mathbf{j}. \tag{2.106}$$

Considering the variation of velocity $\boldsymbol{v} = (u, v, w)$ one can finally write:

$$\begin{aligned}
\mathrm{d}\boldsymbol{v} &= \mathrm{d}\left(u\mathbf{i}\right) + \mathrm{d}\left(v\mathbf{j}\right) + \mathrm{d}\left(w\mathbf{k}\right) \\
&= \mathrm{d}u\,\mathbf{i} + \mathrm{d}v\,\mathbf{j} + \mathrm{d}w\,\mathbf{k} + u\,\mathrm{d}\mathbf{i} + v\,\mathrm{d}\mathbf{j} + w\,\mathrm{d}\mathbf{k} \\
&= \left[\mathrm{d}u + \frac{\mathrm{dX}_\mathbf{i}}{r}\left(w - v\tan\theta\right)\right]\mathbf{i} + \left(\mathrm{d}v + u\frac{\mathrm{dX}_\mathbf{i}\tan\theta}{r} + w\frac{\mathrm{dX}_\mathbf{j}}{r}\right)\mathbf{j} \\
&\quad + \left(\mathrm{d}w + u\frac{\mathrm{dX}_\mathbf{i}}{r} - v\frac{\mathrm{dX}_\mathbf{j}}{r}\right)\mathbf{k}
\end{aligned}$$

The previous equations concerns the effects of a steady spherical reference frame. In the following, the effects of rotating frame will be discussed. Within the LU model, it is important to understand that $t \mapsto \boldsymbol{e}_i\left(t\right)$ is a bounded variation process. The procedure mimics the deterministic one almost verbatim, showing however some differences. For convenience, the translation of the non-inertial system is *a priori* neglected and the same considerations for the deterministic case hold in this case too. The acceleration term can thus be written in its relevant terms as:

$$\mathbb{D}_t^I \boldsymbol{v} = \mathbb{D}_t^R \boldsymbol{v}' + 2\boldsymbol{\Omega} \times \left(\boldsymbol{v}'\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t'\right) + \boldsymbol{\Omega} \times \boldsymbol{\Omega}\,\mathrm{d}t \times \mathbf{X}_t'. \tag{2.107}$$

where $\mathbb{D}_t^R \boldsymbol{v}'$ accounts also for the metric terms introduced in the initial part of this section. Proof of this statement is given in Section 2.9.

2.7.1 Stochastic hydrostatic primitive equations

Primitive equations of motion correspond to a simplification of the equations of motion for a geophysical fluid that is commonly used in ocean science. They are derived from Boussinesq equations (2.96) to (2.98) after the application of three approximations:

— The traditional approximation. All the Coriolis terms that, in the horizontal mo-

mentum equations, involve the vertical velocity are neglected, as well as the metric terms concerning vertical velocities;
— The shallow-fluid approximation. The fluid field has a thickness that is far smaller than the radius of the earth;
— Hydrostatic balance. The vertical acceleration is negligible, and in the vertical momentum equation it can be omitted.

As presented, these three conditions are the same as in the deterministic case. However, the third one must be analysed to understand its meaning under the LU framework. While the first two approximations can be applied verbatim to the stochastic setting, neglecting thus random Coriolis and random metric terms, the application of the third leaves room for a relaxation. Indeed, the vertical momentum equation, after neglecting the acceleration, is

$$[\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t - \boldsymbol{v}_s \,\mathrm{d}t] \cdot \nabla w - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla w)\,\mathrm{d}t - \nabla \cdot \left[\frac{\partial \boldsymbol{\vartheta}}{\partial z}\boldsymbol{\sigma}_t\right]\mathrm{d}t = -\frac{\partial p}{\partial z}\,\mathrm{d}t - \frac{\partial \mathrm{d}p_t^{\sigma}}{\partial z} + b \quad (2.108)$$

and thus the pressure and the stochastic pressure can be defined as:

$$p' = \int_{\eta_b}^{z} \left\{ b + \boldsymbol{v}_s \cdot \nabla w + \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla w) + \nabla \cdot \left[\frac{\partial \boldsymbol{\vartheta}}{\partial z}\boldsymbol{\sigma}_t\right] \right\} \mathrm{d}\zeta, \quad (2.109)$$

$$\mathrm{d}p_t^{\sigma} = \int_{\eta_b}^{z} \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla w \,\mathrm{d}\zeta, \quad (2.110)$$

where w is obtained as usual by integration of the continuity equation

$$w\,(z) = w\,(\eta_b) - \int_{\eta_b}^{z} \left(\nabla_{\mathrm{H}} \cdot \boldsymbol{u}_{\mathrm{H}} - \nabla \cdot \boldsymbol{v}_s\right) \mathrm{d}\zeta \quad (2.111)$$

and η_b is the bottom of the ocean. This type of relaxed hydrostatic balance has been considered in Jamet et al. (2023). The stochastic primitive equations constitute the main focus of this thesis, and they are gathered below.

Stochastic hydrostatic primitive equations

Horizontal momentum: (2.112)

$$\mathbb{D}_t \boldsymbol{u} - \nabla_{\mathrm{H}} \cdot [(\nabla \boldsymbol{\vartheta})\, \boldsymbol{\sigma}_t]\, \mathrm{d}t + f \mathbf{e}_3 \times (\boldsymbol{u}\, \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}) = -\nabla_{\mathrm{H}} (p'\, \mathrm{d}t + \mathrm{d}p_t^{\sigma})$$

Vertical momentum: (2.113)

$$w(z) = w(\eta_b) - \int_{\eta_b}^{z} (\nabla_{\mathrm{H}} \cdot \boldsymbol{u}_{\mathrm{H}} - \nabla \cdot \boldsymbol{v}_s)\, \mathrm{d}\zeta,$$

Hydrostatic pressure and stochastic pressure: (2.114a,b)

$$p'(\boldsymbol{x}) = \int_{\eta_b}^{z} b + \boldsymbol{v}_s \cdot \nabla w + \frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla w) + \nabla \cdot \left[\frac{\partial \boldsymbol{\vartheta}}{\partial z} \boldsymbol{\sigma}_t \right] \mathrm{d}\zeta,$$

$$\mathrm{d}p_t^{\sigma}(\boldsymbol{x}) = \int_{\eta_b}^{z} \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla w\, \mathrm{d}\zeta,$$

Buoyancy transport: (2.115)

$$\mathbb{D}_t \left(b + N^2 z \right) = 0 \tag{2.116}$$

Incompressibility: (2.117a,b)

$$\nabla \cdot [\boldsymbol{v} - \boldsymbol{v}_s] = 0, \qquad \nabla \cdot \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = 0,$$

2.8 Stochastic shallow water equations

The shallow water model conventionally describes the motion of a shallow layer of constant density fluid in hydrostatic balance, bounded from above by a fluid of negligible inertia and from below by a rigid surface (in the single layer model) or a moving free surface of a denser fluid (in the multi-layer model). It is here introduced for its importance in describing the free surface dynamics. From this definition Vallis (2017) one can already infer some properties of the forthcoming Rotating Shallow Water system under Location Uncertainty (RSWLU). First of all, assuming a constant density modifies the stochastic transport operator as explained in Section 2.4.1, thus prescribing the two continuity conditions

$$\nabla \cdot (\boldsymbol{v} - \boldsymbol{v}_s) = 0, \qquad \nabla \cdot \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = 0. \tag{2.118}$$

In particular, this assumption states that the transport velocity in the problem is $\boldsymbol{V} = (\boldsymbol{v} - \boldsymbol{v}_s) + \boldsymbol{\sigma}_t \dot{\mathbf{B}}_t$. Assuming that the layer of fluid is shallow means that it exists a difference

of scales between horizontal and vertical components of the velocity, so a splitting of the two is possible. Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the cartesian reference and let's define $\nabla_{\text{H}} = [\partial_x, \partial_y, 0]$ so that $\nabla = \nabla_{\text{H}} + \partial_z \mathbf{e}_3$. Considering the usual scaling $\boldsymbol{x} = L\hat{\boldsymbol{x}}$ for the horizontal length scale and $z = H\hat{z}$ for the vertical one, one has from incompressibility

$$U\hat{\nabla}_{\text{H}} \cdot \left(\hat{\boldsymbol{u}} - \frac{1}{\Upsilon}\hat{\boldsymbol{u}}_s\right) = W\frac{L}{H}\hat{\partial}_z\left(\hat{w} - \frac{1}{\Upsilon}\hat{w}_s\right), \qquad \hat{\nabla}_{\text{H}} \cdot \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}^{\text{H}}_t = \frac{L}{H}\hat{\partial}_z \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\mathbf{B}}^z_t, \tag{2.119}$$

with $\boldsymbol{v}_s = (\boldsymbol{u}_s, w)$, meaning that the vertical velocity scales as $W = \frac{H}{L}U$. The three dimensional momentum equations can thus be simplified and a horizontal two-dimensional system can be derived. To completely remove the dependence from vertical velocity the continuity condition is integrated from the bottom to the free surface of the fluid layer so that:

$$\begin{aligned}\int_{\eta_b}^{\eta} \nabla \cdot \boldsymbol{v}^\star \,\mathrm{d}z &= \int_{\eta_b}^{\eta} \frac{\partial w^\star}{\partial z}\,\mathrm{d}z + \int_{\eta_b}^{\eta} \nabla_{\text{H}} \cdot \boldsymbol{u}^\star \,\mathrm{d}z \\ &= w^\star|_{\eta_b}^{\eta} + \nabla_{\text{H}} \cdot \int_{\eta_b}^{\eta} \boldsymbol{u}^\star \,\mathrm{d}z - \boldsymbol{u}^\star|_{\eta} \cdot \nabla_{\text{H}}\eta + \boldsymbol{u}^\star|_{\eta_b} \cdot \nabla_{\text{H}}\eta_b \\ &= [w^\star - \boldsymbol{u}^\star \cdot \nabla_{\text{H}} z]_{\eta_b}^{\eta} + \nabla_{\text{H}} \cdot \int_{\eta_b}^{\eta} \boldsymbol{u}^\star \,\mathrm{d}z \end{aligned} \tag{2.120}$$

and for the noise

$$\begin{aligned}\int_{\eta_b}^{\eta} \nabla \cdot \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \,\mathrm{d}z &= \int_{\eta_b}^{\eta} \frac{\partial \sigma_t \mathrm{dB}^z_t}{\partial z}\,\mathrm{d}z + \int_{\eta_b}^{\eta} \nabla_{\text{H}} \cdot \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{\text{H}}_t \,\mathrm{d}z \\ &= [\sigma_t \mathrm{dB}^z_t - \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{\text{H}}_t \cdot \nabla_{\text{H}} z]_{\eta_b}^{\eta} + \nabla_{\text{H}} \cdot \int_{\eta_b}^{\eta} \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{\text{H}}_t \,\mathrm{d}z \end{aligned} \tag{2.121}$$

Summing the two conditions one has

$$\begin{aligned}\int_{\eta_b}^{\eta} \nabla \cdot (\boldsymbol{v}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t)\,\mathrm{d}z = [w^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^z_t - (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{\text{H}}_t) \cdot \nabla(\eta - \eta_b)]_{\eta_b}^{\eta} \\ + \nabla_{\text{H}} \cdot \int_{\eta_b}^{\eta} (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{\text{H}}_t)\,\mathrm{d}z\end{aligned}$$

In the LU formalism the vertical velocity of the two material surfaces $z = \eta$ and $z = \eta_b$ are

$$\mathbb{D}_t \eta = [(w - w_s)\,\mathrm{d}t + \sigma_t \mathrm{dB}^z_t]_{\eta} \tag{2.122}$$

$$\mathbb{D}_t \eta_b = [(w - w_s)\,\mathrm{d}t + \sigma_t \mathrm{dB}^z_t]_{\eta_b} \tag{2.123}$$

so that

$$\int_{\eta_b}^{\eta} \nabla \cdot (\boldsymbol{v}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t)\, \mathrm{d}z = \mathbb{D}_t^{\mathrm{H}} \eta - \mathbb{D}_t^{\mathrm{H}} \eta_b - (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}) \cdot \nabla \eta|_{\eta_b}^{\eta} + \nabla_{\mathrm{H}} \cdot \int_{\eta_b}^{\eta} (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}})\, \mathrm{d}z \tag{2.124}$$

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

$$\mathbb{D}_t^{\mathrm{H}} \theta = d_t \theta + [\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \,\mathrm{d}\mathbf{B}_t^{\mathrm{H}}] \cdot \nabla_{\mathrm{H}} \theta - \frac{1}{2} \nabla_{\mathrm{H}} \cdot (\boldsymbol{a}_{\mathrm{H}} \nabla_{\mathrm{H}} \theta)\, \mathrm{d}t. \tag{2.125}$$

The vertically integrated divergence thus reads

$$\int_{\eta_b}^{\eta} \nabla \cdot (\boldsymbol{v}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t)\, \mathrm{d}z = \mathrm{d}_t (\eta - \eta_b) + \nabla_{\mathrm{H}} \cdot \int_{\eta_b}^{\eta} (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}})\, \mathrm{d}z - \frac{1}{2} \nabla \cdot [\boldsymbol{a}_{\mathrm{H}} \nabla_{\mathrm{H}} (\eta - \eta_b)]\, \mathrm{d}t. \tag{2.126}$$

Since the continuity equation was set to zero, one finds by setting $h = \eta - \eta_b$,

$$\mathrm{d}_t h + \nabla_{\mathrm{H}} \cdot \int_{\eta_b}^{\eta} (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}})\, \mathrm{d}z - \frac{1}{2} \nabla_{\mathrm{H}} \cdot (\boldsymbol{a}_{\mathrm{H}} \nabla_{\mathrm{H}} h)\, \mathrm{d}t = 0 \tag{2.127}$$

The LU model has two additional terms with respect to the original model, which are the horizontal noise, $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}$, and the horizontal diffusion $\frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla h)$. The hydrostatic distribution of pressure presents, in this context reads

$$\frac{1}{\rho_0} \frac{\partial}{\partial z} (p \,\mathrm{d}t + \mathrm{d}p_t^\sigma) = -g \,\mathrm{d}t, \tag{2.128}$$

and it can be decoupled into two equations, the usual relation for hydrostasy and a barotropic characterization of the stochastic pressure, i.e.

$$\frac{\partial p}{\partial z} = -g\rho_0, \tag{2.129}$$

$$\frac{\partial \mathrm{d}p_t^\sigma}{\partial z} = 0. \tag{2.130}$$

Integrating the previous relations from a generic depth z to the sea surface height one has

$$p\,(x, y, z, t) = p\,(x, y, \eta, t) + \int_z^{\eta} \rho g \,\mathrm{d}\varsigma \tag{2.131}$$

and

$$\mathrm{d}p_t^\sigma\left(x,y,z,t\right)=\mathrm{d}p_t^\sigma\left(x,y,\eta,t\right). \tag{2.132}$$

Neglecting the influence of atmospheric pressure, that is setting $p\left(x,y,\eta,t\right)=0$, and splitting the pressure into a surface deviation component, p_s, and a hydrostatic component, p_h, one has

$$p=\int_0^\eta \rho g\,\mathrm{d}\varsigma+\int_z^0 \rho g\,\mathrm{d}\varsigma=\rho g\eta-\rho g z. \tag{2.133}$$

In the horizontal momentum equations pressure acts in terms of its horizontal gradient, so it comes without saying that the only component of interest is $\rho g\eta$, as it is the only one depending on the two coordinates x and y. This leads to the horizontal momentum equation, after the substitution of the material derivative with its stochastic counterpart,

$$\mathbb{D}_t^{\mathrm{H}}\boldsymbol{u}+f\mathbf{e}_3\times\left(\boldsymbol{u}\,\mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right)-\nabla_{\mathrm{H}}\cdot\left[\left(\nabla\boldsymbol{\vartheta}\right)\boldsymbol{\sigma}_t\right]\mathrm{d}t=-\frac{1}{\rho_0}\rho g\nabla_{\mathrm{H}}\eta\,\mathrm{d}t-\frac{1}{\rho_0}\nabla_{\mathrm{H}}\mathrm{d}p_t^\sigma. \tag{2.134}$$

The stochastic rotating shallow water equations are summarized in the box below.

Stochastic rotating shallow water equations

Horizontal momentum: (2.135)

$$\mathbb{D}_t^{\mathrm{H}}\boldsymbol{u}+f\mathbf{e}_3\times\left(\boldsymbol{u}\,\mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right)-\nabla_{\mathrm{H}}\cdot\left[\left(\nabla\boldsymbol{\vartheta}\right)\boldsymbol{\sigma}_t\right]\mathrm{d}t=-\frac{1}{\rho_0}\rho g\nabla_{\mathrm{H}}\eta\,\mathrm{d}t-\frac{1}{\rho_0}\nabla_{\mathrm{H}}\mathrm{d}p_t^\sigma$$

Mass conservation: (2.136)

$$\mathrm{d}_t h+\nabla_{\mathrm{H}}\cdot\int_{\eta_b}^{\eta}\left(\boldsymbol{u}^\star\,\mathrm{d}t+\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right)\,\mathrm{d}z-\frac{1}{2}\nabla_{\mathrm{H}}\cdot\left(\boldsymbol{a}_{\mathrm{H}}\nabla_{\mathrm{H}}h\right)\,\mathrm{d}t=0$$

In particular, this version includes the covariation effects of the random pressure on the smooth-in-time component of the flow. In other works, such as Brecht et al., 2021, this term was neglected and so is done in the rest of this document.

Barotropic flows

Shallow water equations are usually simpler than the one previously introduced, because they rely on a barotropic assumption, that is assuming $\partial_z\boldsymbol{u}=0$. This assumption can be made also for the RSWLU system. Considering the stochastic pressure terms, it

is clear that

$$\frac{\partial^2 p}{\partial x \partial z} = 0, \qquad \frac{\partial^2 \left(\mathrm{d}p_t^\sigma\right)}{\partial x \partial z} = 0,$$

so taking the vertical derivative of the horizontal momentum equations

$$\frac{\partial}{\partial z}\left(\mathbb{D}_t^{\mathrm{H}} \boldsymbol{u}\right) + f\mathbf{e}_3 \times \left(\partial_z \boldsymbol{u}\,\mathrm{d}t + \partial_z \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right) = 0 \tag{2.137}$$

and reversing the order of differentiation

$$\mathbb{D}_t\left(\partial_z \boldsymbol{u}\right) + \left[\partial_z \boldsymbol{u}\,\mathrm{d}t + \partial_z \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right] \cdot \nabla_{\mathrm{H}} \boldsymbol{u} + f\mathbf{e}_3 \times \left(\partial_z \boldsymbol{u}\,\mathrm{d}t + \partial_z \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right) = 0 \tag{2.138}$$

it follows that if $\partial_z \boldsymbol{u} = 0$ and $\partial_z \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}} = 0$ at any time, then this property is conserved and it makes sense to assume that the horizontal velocity has no vertical variation with respect to the LU transport operator. Having established that the two conditions $\partial_z \boldsymbol{u} = 0$ and $\partial_z \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}} = 0$ are reasonable, one can write a modified continuity equation and represent the RSWLU as

$$\mathbb{D}_t \boldsymbol{u} + f\mathbf{e}_3 \times \left(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right) = -\frac{1}{\rho_0}\rho g \nabla_{\mathrm{H}} \eta\,\mathrm{d}t - \frac{1}{\rho_0}\nabla_{\mathrm{H}} \mathrm{d}p_t^\sigma \tag{2.139}$$

$$\mathrm{d}_t h + \nabla_{\mathrm{H}} \cdot \left[h\left(\boldsymbol{u}^\star\,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{\mathrm{H}}\right)\right] - \frac{1}{2}\nabla_{\mathrm{H}} \cdot \left(\boldsymbol{a}_{\mathrm{H}} \nabla_{\mathrm{H}} h\right)\,\mathrm{d}t = 0. \tag{2.140}$$

Conclusion

In this chapter, the fundamental models for geophysical fluid mechanics under location uncertainty were derived. All the covariation terms between the stochastic forcing and the stochastic transport were retained, contrarily to other presentation in the literature. Some of terms can be assumed to be negligible on the base of magnitude analysis such as in Tissot, Mémin, et al. (2023). It is fundamental to remark that these terms appear because the underlying Brownian term is the same for both the stochastic transport and the stochastic forcing, thus describing the effects of the forcing on the non resolved scales. In case two different Brownian motions are chosen for the stochastic transport and the stochastic forcing, all the covariation terms would disappear due to the independence of the two Brownian terms.

2.9 Detailed proofs of the previous statements

The path followed is the one outlined in Cintolesi et al. (2020b), that carries out the original derivation of Mémin (2014), made explicit with some of the calculations from Bauer et al. (2020b).

Proof of validity of Equation 2.14. The transported process $\theta(\mathbf{X}_t, t)$ is a composition of two processes, as it is the semi-martingale $\theta(\boldsymbol{x}, t)$ represented in Equation (2.13), composed with the Lagrangian stochastic flow $\mathbf{X}_t$ of Equation (2.6). To differentiate in time the total process, the *Itô-Wentzell formula* (see Kunita, 1997, Theorem 3.8) is needed:

$$\mathrm{d}\theta(\mathbf{X}_t, t) = \mathrm{d}_t\theta(\mathbf{X}_t, t) + \frac{\partial\theta(\mathbf{X}_t, t)}{\partial x_i}\,\mathrm{dX}_t^i + \frac{1}{2}\frac{\partial^2\theta(\mathbf{X}_t, t)}{\partial x_i \partial x_j}\,\mathrm{d}\left\langle \mathrm{X}^i, \mathrm{X}^j\right\rangle_t + \mathrm{d}\left\langle \frac{\partial\theta(\mathbf{X}, \cdot)}{\partial x_i}, \mathrm{X}^i\right\rangle_t. \tag{2.141}$$

Since that $\int_0^t g(\mathbf{x}, s)\,\mathrm{d}s$ is a *bounded variation process* then the brackets (i.e. the quadratic variation) can be computed as

$$\begin{aligned}
\left\langle \mathrm{X}^i, \mathrm{X}^j\right\rangle_t &= \left\langle \int_0^{\cdot}\int_\Omega \breve{\sigma}_{ik}(\mathbf{X}_s, \boldsymbol{y}, s)\,\mathrm{dB}_s^k(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \int_0^{\cdot}\int_\Omega \breve{\sigma}_{j\ell}(\mathbf{X}_s, \boldsymbol{z}, s)\,\mathrm{dB}_s^\ell(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}\right\rangle_t \\
&= \int_0^t \int_{\Omega\times\Omega} \breve{\sigma}_{ik}(\mathbf{X}_s, \boldsymbol{y}, s)\,\breve{\sigma}_{j\ell}(\mathbf{X}_s, \boldsymbol{z}, s)\,\underbrace{\mathrm{d}\left\langle \mathrm{B}^k(\boldsymbol{y}), \mathrm{B}^\ell(\boldsymbol{z})\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s}\,\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \\
&= \int_0^t \int_\Omega \breve{\sigma}_{ik}(\mathbf{X}_s, \boldsymbol{y}, s)\,\breve{\sigma}_{kj}(\mathbf{X}_s, \boldsymbol{y}, s)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}s \qquad \forall i, j = 1, \ldots, d.
\end{aligned}$$

Concluding this calculation, we can state

$$\frac{\mathrm{d}\left\langle \mathrm{X}^i, X^j\right\rangle_t}{\mathrm{d}t} = \int_\Omega \breve{\sigma}_{ik}(\mathbf{X}_t, \boldsymbol{y}, t)\,\breve{\sigma}_{kj}(\mathbf{X}_t, \boldsymbol{y}, t)\,\mathrm{d}\boldsymbol{y} = a_{ij}(\mathbf{X}_t, t), \qquad \forall i, j = 1, \ldots, d. \tag{2.142}$$

Then

$$\frac{1}{2}\frac{\partial^2\theta(\mathbf{X}_t, t)}{\partial x_i \partial x_j}\,\mathrm{d}\left\langle \mathrm{X}^i, \mathrm{X}^j\right\rangle_t = \frac{1}{2}a_{ij}(\mathbf{X}_t, t)\,\frac{\partial^2\theta(\mathbf{X}_t, t)}{\partial x_i \partial x_j}\,\mathrm{d}t. \tag{2.143}$$

The fourth term in (2.141) can be obtained first differentiating (2.13) in x_i and later on computing the quadratic variation.

$$\frac{\partial\theta(\boldsymbol{x}, t)}{\partial x_i} = \frac{\partial\theta(\boldsymbol{x}, 0)}{\partial x_i} + \int_0^t \frac{\partial g(\boldsymbol{x}, s)}{\partial x_i}\,\mathrm{d}s + \int_0^t\int_\Omega \frac{\partial f_k(\boldsymbol{x}, \boldsymbol{y}, s)}{\partial x_i}\,\mathrm{dB}_s^k(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}s, \tag{2.144}$$

resulting in

$$\mathrm{d}\left\langle \frac{\partial\theta}{\partial x_i}, \mathrm{X}^i \right\rangle_t = \int_\Omega \breve{\sigma}_{ij}\left(\mathbf{X}_t, \boldsymbol{y}, t\right) \frac{\partial f_j}{\partial x_i}\left(\mathbf{X}_t, \boldsymbol{y}, t\right) \mathrm{d}\boldsymbol{y}\, \mathrm{d}t. \tag{2.145}$$

The equation for the differential of a transported process (2.14) is thus proved. ■

Proof of validity of equations 2.15 and 2.16. Starting from the differential of the transported process, that is Equation (2.14), expressing the path of the stochastic flow as an Itô diffusion process (2.5), the following formula is obtained:

$$\mathrm{d}\theta = \mathrm{d}_t\theta + \frac{\partial\theta}{\partial x_i}\left(v_i\,\mathrm{d}t + \int_\Omega \breve{\sigma}_{ik}\mathrm{dB}_t^k \mathrm{d}\boldsymbol{y}\right) + \frac{1}{2}a_{ij}\frac{\partial^2\theta}{\partial x_i\partial x_j}\,\mathrm{d}t + \int_\Omega \breve{\sigma}_{ij}\frac{\partial}{\partial x_i}f_j\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t, \tag{2.146}$$

Expanding the terms in the round brackets,

$$\mathrm{d}\theta = \mathrm{d}_t\theta + \frac{\partial\theta}{\partial x_i}v_i\,\mathrm{d}t + \int_\Omega \frac{\partial\theta}{\partial x_i}\breve{\sigma}_{ik}\mathrm{dB}_t^k \mathrm{d}\boldsymbol{y} + \frac{1}{2}a_{ij}\frac{\partial^2\theta}{\partial x_i\partial x_j}\,\mathrm{d}t + \int_\Omega \breve{\sigma}_{ij}\frac{\partial}{\partial x_i}f_j\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t. \tag{2.147}$$

Imposing conservation of the quantity $\theta\left(\mathbf{X}_t, t\right)$, that is $\mathrm{d}\theta\left(\mathbf{X}_t, t\right) \overset{!}{=} 0$, one finds

$$\mathrm{d}_t\theta = -\frac{\partial\theta}{\partial x_i}v_i\,\mathrm{d}t - \frac{1}{2}a_{ij}\frac{\partial^2\theta}{\partial x_i\partial x_j}\,\mathrm{d}t - \int_\Omega \breve{\sigma}_{ij}\frac{\partial}{\partial x_i}f_j\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t - \int_\Omega \frac{\partial\theta}{\partial x_i}\breve{\sigma}_{ik}\mathrm{dB}_t^k\,\mathrm{d}\boldsymbol{y} \tag{2.148}$$

thus one can identify g and f in Equation (2.13) (as this decomposition is proved to be unique) as

$$\begin{aligned} g\left(\mathbf{X}_t, t\right) &= -\frac{\partial\theta\left(\mathbf{X}_t, t\right)}{\partial x_i}v_i\left(\mathbf{X}_t, t\right) - \frac{1}{2}a_{ij}\left(\mathbf{X}_t, t\right)\frac{\partial^2\theta\left(\mathbf{X}_t, t\right)}{\partial x_i\partial x_j} \\ &\qquad - \int_\Omega \breve{\sigma}_{ij}\left(\mathbf{X}_t, \boldsymbol{y}, t\right)\frac{\partial}{\partial x_i}f_j\left(\mathbf{X}_t, \boldsymbol{y}, t\right)\mathrm{d}\boldsymbol{y} \\ f_j\left(\mathbf{X}_t, \boldsymbol{y}, t\right) &= -\frac{\partial\theta\left(\mathbf{X}_t, t\right)}{\partial x_k}\breve{\sigma}_{kj}\left(\mathbf{X}_t, \boldsymbol{y}, t\right). \end{aligned}$$

The last equality is recovered through the relation

$$\int_\Omega \left[f_k\left(\mathbf{X}_t, \boldsymbol{y}, t\right) + \frac{\partial\theta\left(\mathbf{X}_t, t\right)}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_t, \boldsymbol{y}, t\right)\right]\mathrm{dB}_t^k\left(\boldsymbol{y}\right)\mathrm{d}\boldsymbol{y} = 0 \tag{2.149}$$

that is valid for every choice of the Brownian motion dB_t^k. Inserting the expression for f

in the last integral in g, one can carry out

$$\begin{aligned}
\int_\Omega \breve{\sigma}_{ij}\frac{\partial f_j}{\partial x_i}\,\mathrm{d}\boldsymbol{y} &= -\frac{\partial^2\theta}{\partial x_i\partial x_k}\underbrace{\int_\Omega \breve{\sigma}_{ij}\breve{\sigma}_{jk}\,\mathrm{d}\boldsymbol{y}}_{a_{ik}(\mathbf{X}_t,t)} - \frac{\partial\theta}{\partial x_k}\int_\Omega \breve{\sigma}_{ij}\frac{\partial\breve{\sigma}_{jk}}{\partial x_i}\,\mathrm{d}\boldsymbol{y} \\
&= -a_{ik}\frac{\partial^2\theta}{\partial x_i\partial x_k} - \frac{\partial\theta}{\partial x_k}\int_\Omega\left[\frac{\partial}{\partial x_i}\left(\breve{\sigma}_{ij}\breve{\sigma}_{jk}\right) - \frac{\partial\breve{\sigma}_{ij}}{\partial x_i}\breve{\sigma}_{jk}\right]\mathrm{d}\boldsymbol{y} \\
&= -a_{ik}\frac{\partial^2\theta}{\partial x_i\partial x_k} - \frac{\partial\theta}{\partial x_k}\frac{\partial a_{ik}}{\partial x_i} + \frac{\partial\theta}{\partial x_k}\int_\Omega \breve{\sigma}_{jk}\frac{\partial\breve{\sigma}_{ij}}{\partial x_i}\,\mathrm{d}\boldsymbol{y} \\
&= -a_{ik}\frac{\partial^2\theta}{\partial x_i\partial x_k} - \frac{1}{2}\frac{\partial\theta}{\partial x_k}\frac{\partial a_{ik}}{\partial x_i} - \frac{1}{2}\frac{\partial\theta}{\partial x_k}\frac{\partial a_{ik}}{\partial x_i} + \frac{\partial\theta}{\partial x_k}\int_\Omega \breve{\sigma}_{jk}\frac{\partial\breve{\sigma}_{ij}}{\partial x_i}\,\mathrm{d}\boldsymbol{y} \\
&= -a_{ik}\frac{\partial^2\theta}{\partial x_i\partial x_k} - \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial\theta}{\partial x_k}\right) + \frac{1}{2}a_{ik}\frac{\partial^2\theta}{\partial x_i\partial x_k} \\
&\qquad - \frac{1}{2}\frac{\partial\theta}{\partial x_k}\frac{\partial a_{ik}}{\partial x_i} + \frac{\partial\theta}{\partial x_k}\int_\Omega \breve{\sigma}_{jk}\frac{\partial\breve{\sigma}_{ij}}{\partial x_i}\,\mathrm{d}\boldsymbol{y} \\
&= -\frac{1}{2}a_{ik}\frac{\partial^2\theta}{\partial x_i\partial x_k} - \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial\theta}{\partial x_k}\right) - \frac{1}{2}\frac{\partial a_{ik}}{\partial x_i}\frac{\partial\theta}{\partial x_k} + \frac{\partial\theta}{\partial x_k}\int_\Omega \breve{\sigma}_{jk}\frac{\partial\breve{\sigma}_{ij}}{\partial x_i}\,\mathrm{d}\boldsymbol{y}.
\end{aligned}$$

The final equation, inserting the previous calculation into (2.148), reads

$$\begin{aligned}
\mathrm{d}_t\theta = -\left[\left(v_k - \frac{1}{2}\frac{\partial a_{ik}}{\partial x_i} + \int_\Omega \breve{\sigma}_{jk}\frac{\partial\breve{\sigma}_{ij}}{\partial x_i}\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}t + \int_\Omega \breve{\sigma}_{kj}\mathrm{dB}_t^j\,\mathrm{d}\boldsymbol{y}\right]\frac{\partial\theta}{\partial x_k} \\
+ \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial\theta}{\partial x_k}\right)\mathrm{d}t.
\end{aligned}$$

■

In the previous proof, all the dependencies were kept to illustrate how the process is characterized. In the following, unless differently specified, every component has to be intended as $\theta = \theta(\boldsymbol{x},t)$, $v_i = v_i(\boldsymbol{x},t)$, $a_{ij} = a_{ij}(\boldsymbol{x},t)$, while the integration kernel reads $\breve{\sigma}_{ij} = \breve{\sigma}_{ij}(\boldsymbol{x},\boldsymbol{y},t)$ and $f_j = f_j(\boldsymbol{x},\boldsymbol{y},t)$. Whenever the process is said to be transported, the mapping $\boldsymbol{x} \to \mathbf{X}_t$ is assumed.

2.9.1 Proof of stochastic transport theorem

Let V_t be an arbitrary control volume contained in Ω, and let $\varphi(\boldsymbol{x}, t)$ be a test function defined over Ω such that it has a representation as a semi-martingale that reads

$$\varphi(\boldsymbol{x}, t) = \varphi(\boldsymbol{x}, 0) + \int_0^t g(\boldsymbol{x}, s)\, \mathrm{d}s + \int_0^t \int_\Omega f_j(\boldsymbol{x}, \boldsymbol{y}, s)\, \mathrm{dB}_s^j(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}\, \mathrm{d}s. \tag{2.150}$$

Assume furthermore that the function φ has a compact support over V_t and vanishing normal derivative on the boundary ∂V_t (this corresponds to the assumption of φ not diffusing outside of V_t). The process φ is transported by the stochastic flow $\boldsymbol{x}_0 \mapsto \boldsymbol{x} = \mathbf{X}_t(\boldsymbol{x}_0)$, with

$$\mathbf{X}_t(\boldsymbol{x}) = \mathbf{X}_0 + \int_0^t \boldsymbol{v}(\mathbf{X}_s, s)\, \mathrm{d}s + \int_0^t \boldsymbol{\sigma}_s(\mathbf{X}_s)\, \mathrm{d}\mathbf{B}_s, \tag{2.151}$$

starting from an initial condition g

$$\varphi(\mathbf{X}_t(\boldsymbol{x}_0), t) = g(\boldsymbol{x}_0). \tag{2.152}$$

Under these conditions, as proven for Equations 2.15 and 2.16, φ has a time increment defined as

$$\mathrm{d}_t\varphi = -\left[(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}_t^{\mathrm{T}}(\nabla \cdot \boldsymbol{\sigma}_t))\, \mathrm{d}t + \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t\right] \cdot \nabla\varphi + \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla\varphi)\, \mathrm{d}t, \tag{2.153}$$

that in indexes notation reads

$$\mathrm{d}_t\varphi = -\left[\left(v_k - \frac{1}{2}\frac{\partial a_{ik}}{\partial x_i} + \int_\Omega \breve{\sigma}_{jk}\frac{\partial \breve{\sigma}_{ij}}{\partial x_i}\, \mathrm{d}\boldsymbol{y}\right)\mathrm{d}t + \int_\Omega \breve{\sigma}_{kj}\mathrm{dB}_t^j\, \mathrm{d}\boldsymbol{y}\right]\frac{\partial\varphi}{\partial x_k} + \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{ik}\frac{\partial\varphi}{\partial x_k}\right)\mathrm{d}t.$$

This setting corresponds to stating that the the material derivative of φ along the flow is

$$(\mathrm{D}_t\varphi)(\mathbf{X}_t(\boldsymbol{x}_0), t) = \mathrm{d}_t(\varphi(\mathbf{X}_t(\boldsymbol{x}_0), t)) = \mathrm{d}_t g(\boldsymbol{x}_0). \tag{2.154}$$

The temporal increment of φ is split into its two main components, the bounded variation and the martingale term, as

$$\mathrm{d}_t\varphi = -\mathcal{L}\varphi\, \mathrm{d}t - \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t \cdot \nabla\varphi. \tag{2.155}$$

with

$$\mathcal{L}\varphi = \boldsymbol{v}^\star \cdot \nabla\varphi - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla\varphi). \tag{2.156}$$

At this point, it is worth introducing the adjoint operator $\mathcal{L}^*$ of $\mathcal{L}$, that is the operator such that $(\mathcal{L}\varphi, \psi) = (\varphi, \mathcal{L}^*\psi)$, where (f, g) denotes the inner product between f and g. The full definition is

$$\int_\Omega \left[\boldsymbol{v}^\star \cdot \nabla\varphi - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla\varphi)\right] \psi \,\mathrm{d}\boldsymbol{x} = \int_\Omega \varphi \left[-\nabla \cdot (\boldsymbol{v}^\star\psi) - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla\psi)\right] \,\mathrm{d}\boldsymbol{x}, \tag{2.157}$$

as a result of *Green's lemma* on the first term

$$\int_\Omega v_k^\star \psi \frac{\partial\varphi}{\partial x_k} \,\mathrm{d}\boldsymbol{x} = -\int_\Omega \varphi \frac{\partial\left(v_k^\star\psi\right)}{\partial x_k} \,\mathrm{d}\boldsymbol{x} + \underbrace{\int_{\partial\Omega} \varphi \left(v_k^\star\psi\right) n^k \,\mathrm{d}\varsigma}_{\varphi|_{\partial\Omega}=0}, \tag{2.158}$$

and repeated application of Green's lemma and exploitation of the symmetry of $\boldsymbol{a}$ in the second term

$$\begin{aligned}\int_\Omega \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial\varphi}{\partial x_j}\right)\psi \,\mathrm{d}\boldsymbol{x} &= \int_{\partial\Omega} \psi a_{ij}\frac{\partial\varphi}{\partial x_j} n_i \,\mathrm{d}\varsigma - \int_\Omega a_{ij}\frac{\partial\varphi}{\partial x_j}\frac{\partial\psi}{\partial x_i} \,\mathrm{d}\boldsymbol{x} \\ &= \underbrace{\int_{\partial\Omega} \psi a_{ij}\frac{\partial\varphi}{\partial x_j} n_i \,\mathrm{d}\varsigma}_{\boldsymbol{a}\nabla\varphi\cdot\boldsymbol{n}|_{\partial\Omega}=0} - \underbrace{\int_{\partial\Omega} \varphi a_{ji}\frac{\partial\psi}{\partial x_i} n_j \,\mathrm{d}\varsigma}_{\varphi|_{\partial\Omega}=0} + \int_\Omega \frac{\partial}{\partial x_j}\left(a_{ji}\frac{\partial\psi}{\partial x_i}\right)\varphi \,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Introduce now the process q that follows a *weak balance equation* defined as

$$\mathrm{d}\int_{V_t} \varphi q \,\mathrm{d}\boldsymbol{x} = \int_{V_t} \varphi \left(Q_t \,\mathrm{d}t + \boldsymbol{Q}_\sigma \cdot \mathrm{d}\mathbf{B}_t\right) \,\mathrm{d}\boldsymbol{x}, \tag{2.159}$$

with $Q_t \,\mathrm{d}t$ the bounded variation component of the forcing of q and $\boldsymbol{Q}_\sigma \cdot \mathrm{d}\mathbf{B}_t$ the martingale term of q. The left-hand side can be developed following the product rule and Itô's formula as

$$\begin{aligned}\mathrm{d}\int_{V_t} \varphi q \,\mathrm{d}\boldsymbol{x} &= \mathrm{d}\int_\Omega \varphi q \,\mathrm{d}\boldsymbol{x} \\ &= \int_\Omega \left[q \,\mathrm{d}_t\varphi + \varphi \,\mathrm{d}_t q + \mathrm{d}\left\langle\varphi, q\right\rangle_t\right] \,\mathrm{d}\boldsymbol{x} \\ &= \int_\Omega \left[q \left(-\mathcal{L}\varphi \,\mathrm{d}t - \boldsymbol{\sigma}_t \,\mathrm{d}\mathbf{B}_t \cdot \nabla\varphi\right) + \varphi \,\mathrm{d}_t q + \mathrm{d}\left\langle\varphi, q\right\rangle_t\right] \,\mathrm{d}\boldsymbol{x} \\ &= \int_\Omega \left\{\varphi\left[\mathrm{d}_t q - \mathcal{L}^* q \,\mathrm{d}t + \nabla \cdot (q\boldsymbol{\sigma}_t \,\mathrm{d}\mathbf{B}_t)\right] + \mathrm{d}\left\langle\varphi, q\right\rangle_t\right\} \,\mathrm{d}\boldsymbol{x} \\ &= \int_\Omega \left\{\varphi\left[\mathrm{d}_t q + \nabla \cdot (\boldsymbol{v}^\star q) \,\mathrm{d}t + \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla q) \,\mathrm{d}t + \nabla \cdot (q\boldsymbol{\sigma}_t \,\mathrm{d}\mathbf{B}_t)\right] + \mathrm{d}\left\langle\varphi, q\right\rangle_t\right\} \,\mathrm{d}\boldsymbol{x},\end{aligned}$$

so that the balance of processes acting reads thus

$$\int_\Omega \Big\{ \varphi \Big[\mathrm{d}_t q + \nabla \cdot (\boldsymbol{v}^\star q)\, \mathrm{d}t + \frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla q)\, \mathrm{d}t + \nabla \cdot (q \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t) - Q_t\, \mathrm{d}t - \boldsymbol{Q}_\sigma \cdot \mathrm{d}\mathbf{B}_t \Big] + \mathrm{d} \langle \varphi, q \rangle_t \Big\}\, \mathrm{d}\boldsymbol{x} = 0.$$

At this stage, the quadratic covariation $\mathrm{d} \langle \varphi, q \rangle_t$ has to be defined. The dynamics of the process q is not yet fully defined, as the term $\mathrm{d} \langle \varphi, q \rangle_t$ is missing, but its martingale part is already completely defined as the quadratic covariation is going to provide a bounded variation term. The dynamics of q can thus be defined as

$$\mathrm{d}_t q = f\, \mathrm{d}t + (\boldsymbol{Q}_\sigma \cdot \mathrm{d}\mathbf{B}_t - \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t \cdot \nabla q - q \nabla \cdot \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t), \tag{2.160}$$

where f encloses all the bounded variation terms of $\mathrm{d}_t q$, still to be completely defined but that will not play a role in the definition of $\mathrm{d} \langle \varphi, q \rangle_t$. The quadratic covariation can then be computed and the result is:

$$\begin{aligned}
\int_\Omega \mathrm{d}\langle q, \varphi \rangle_t\, \mathrm{d}\boldsymbol{x} = &- \int_\Omega \mathrm{d}\langle \boldsymbol{Q}_\sigma \cdot \mathrm{d}\mathbf{B}_t, \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle_t\, \mathrm{d}\boldsymbol{x} && (2.161)\\
&+ \int_\Omega \mathrm{d}\langle \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t \cdot \nabla q, \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle_t\, \mathrm{d}\boldsymbol{x} && (2.162)\\
&+ \int_\Omega \mathrm{d}\langle q \nabla \cdot \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t, \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle_t\, \mathrm{d}\boldsymbol{x} && (2.163)\\
= &+ \int_\Omega \varphi \left[\nabla \cdot \boldsymbol{\sigma}_t \boldsymbol{Q}_\sigma \right] \mathrm{d}\boldsymbol{x}\, \mathrm{d}t && (2.164)\\
&- \int_\Omega \varphi \left[\nabla \cdot (\boldsymbol{a} \nabla q) \right] \mathrm{d}\boldsymbol{x}\, \mathrm{d}t && (2.165)\\
&- \int_\Omega \varphi \left[\nabla \cdot (q \boldsymbol{\sigma}_t^{\mathrm{T}} (\nabla \cdot \boldsymbol{\sigma}_t)) \right] \mathrm{d}\boldsymbol{x}\, \mathrm{d}t. && (2.166)
\end{aligned}$$

The balance equation becomes

$$\int_\Omega \varphi \Big[\mathrm{d}_t q + \nabla \cdot [q (\boldsymbol{v} - \boldsymbol{v}^s)\, \mathrm{d}t + q \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t] - \frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla q)\, \mathrm{d}t + \nabla \cdot \boldsymbol{\sigma}_t \boldsymbol{Q}_\sigma\, \mathrm{d}t - Q_t\, \mathrm{d}t - \boldsymbol{Q}_\sigma \cdot \mathrm{d}\mathbf{B}_t \Big]\, \mathrm{d}\boldsymbol{x} = 0.$$

Identifying φ as the characteristic function of the volume V_t, that is $\varphi = \mathbb{1}(V_t)$, one reaches the final result:

$$\int_{V_t} \Big[\mathrm{d}_t q + \nabla \cdot [q (\boldsymbol{v} - \boldsymbol{v}^s)\, \mathrm{d}t + q \boldsymbol{\sigma}_t\, \mathrm{d}\mathbf{B}_t] - \frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla q)\, \mathrm{d}t$$

$$+\nabla\cdot\boldsymbol{\sigma}_t\boldsymbol{Q}_\sigma\,\mathrm{d}t\Big]\,\mathrm{d}\boldsymbol{x}=\int_{V_t}\left[Q_t\,\mathrm{d}t+\boldsymbol{Q}_\sigma\cdot\mathrm{d}\mathbf{B}_t\right]\mathrm{d}\boldsymbol{x}.$$

Proof of equivalence between integral 2.161 and 2.164.

$$\begin{aligned}\left\langle\boldsymbol{Q}_\sigma\cdot\mathrm{d}\mathbf{B}_t,\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\varphi\right\rangle&\overset{\text{def}}{=}\mathrm{d}\left\langle\int_0^{\cdot}\int_\Omega\breve{Q}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{dB}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y},\int_0^{\cdot}\int_\Omega\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{j\ell}\left(\mathbf{X}_s,\boldsymbol{z},s\right)\,\mathrm{dB}_s^\ell\left(\boldsymbol{z}\right)\,\mathrm{d}\boldsymbol{z}\right\rangle_t\\&=\mathrm{d}\int_0^t\int_{\Omega\times\Omega}\breve{Q}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{j\ell}\left(\mathbf{X}_s,\boldsymbol{z},s\right)\underbrace{\mathrm{d}\left\langle\mathrm{B}^k\left(\boldsymbol{y}\right),\mathrm{B}^\ell\left(\boldsymbol{z}\right)\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s}\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z}\\&=\mathrm{d}\int_0^t\int_\Omega\breve{Q}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}s\\&=\int_\Omega\breve{Q}_\sigma^k\left(\mathbf{X}_t,\boldsymbol{y},t\right)\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}t\end{aligned}$$

Integrating by parts

$$\begin{aligned}-\int_\Omega\left\langle\boldsymbol{Q}_\sigma\cdot\mathrm{d}\mathbf{B}_t,\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\varphi\right\rangle\mathrm{d}\boldsymbol{x}&=-\int_\Omega\int_\Omega\breve{Q}_\sigma^k\left(\mathbf{X}_t,\boldsymbol{y},t\right)\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\&=-\int_\Omega\frac{\partial\varphi}{\partial x_j}\int_\Omega\breve{Q}_\sigma^k\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\&=\int_\Omega\varphi\frac{\partial}{\partial x_j}\int_\Omega\breve{Q}_\sigma^k\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\&=\int_\Omega\varphi\left[\nabla\cdot\boldsymbol{\sigma}_t\boldsymbol{Q}_\sigma\right]\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\end{aligned}$$

■

Proof of equivalence between integral 2.162 and 2.165.

$$\begin{aligned}\left\langle\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla q,\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\varphi\right\rangle&\overset{\text{def}}{=}\mathrm{d}\left\langle\int_0^{\cdot}\int_\Omega\frac{\partial q}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{dB}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y},\int_0^{\cdot}\int_\Omega\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{j\ell}\left(\mathbf{X}_s,\boldsymbol{z},s\right)\,\mathrm{dB}_s^\ell\left(\boldsymbol{z}\right)\,\mathrm{d}\boldsymbol{z}\right\rangle_t\\&=\mathrm{d}\int_0^t\int_{\Omega\times\Omega}\frac{\partial q}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{j\ell}\left(\mathbf{X}_s,\boldsymbol{z},s\right)\underbrace{\mathrm{d}\left\langle\mathrm{B}^k\left(\boldsymbol{y}\right),\mathrm{B}^\ell\left(\boldsymbol{z}\right)\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s}\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z}\\&=\mathrm{d}\int_0^t\int_\Omega\frac{\partial q}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial\varphi}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}s\\&=\frac{\partial q}{\partial x_i}\frac{\partial\varphi}{\partial x_j}\int_\Omega\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t\end{aligned}$$

Integrating by parts one has

$$\begin{aligned}\int_\Omega\left\langle\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla q,\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\varphi\right\rangle\mathrm{d}\boldsymbol{x}&=\int_\Omega\frac{\partial q}{\partial x_i}\frac{\partial\varphi}{\partial x_j}\int_\Omega\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\&=-\int_\Omega\varphi\frac{\partial}{\partial x_j}\left[\frac{\partial q}{\partial x_i}\int_\Omega\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\right]\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\&=-\int_\Omega\varphi\frac{\partial}{\partial x_j}\left[\frac{\partial q}{\partial x_i}\int_\Omega\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\right]\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\end{aligned}$$

$$= -\int_\Omega \varphi \nabla \cdot [\boldsymbol{a} \nabla q] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{x}$$

■

Proof of equivalence between integral 2.163 and 2.166.

$$\begin{aligned}
\langle q \nabla \cdot \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t, \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle &\overset{\text{def}}{=} \mathrm{d} \left\langle \int_0^{\cdot} \int_\Omega q \frac{\partial}{\partial x_i} \breve{\sigma}_{ik} \left(\mathbf{X}_s, \boldsymbol{y}, s \right) \, \mathrm{d}\mathrm{B}_s^k \left(\boldsymbol{y} \right) \, \mathrm{d}\boldsymbol{y}, \int_0^{\cdot} \int_\Omega \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{j\ell} \left(\mathbf{X}_s, \boldsymbol{z}, s \right) \, \mathrm{d}\mathrm{B}_s^\ell \left(\boldsymbol{z} \right) \, \mathrm{d}\boldsymbol{z} \right\rangle_t \\
&= \mathrm{d} \int_0^t \int_{\Omega \times \Omega} q \frac{\partial}{\partial x_i} \breve{\sigma}_{ik} \left(\mathbf{X}_s, \boldsymbol{y}, s \right) \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{j\ell} \left(\mathbf{X}_s, \boldsymbol{z}, s \right) \underbrace{\mathrm{d} \left\langle \mathrm{B}^k \left(\boldsymbol{y} \right), \mathrm{B}^\ell \left(\boldsymbol{z} \right) \right\rangle_s}_{\delta_{k\ell} \delta(\boldsymbol{y} - \boldsymbol{z}) \, \mathrm{d}s} \mathrm{d}\boldsymbol{y} \mathrm{d}\boldsymbol{z} \\
&= \mathrm{d} \int_0^t \int_\Omega q \frac{\partial}{\partial x_i} \breve{\sigma}_{ik} \left(\mathbf{X}_s, \boldsymbol{y}, s \right) \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{jk} \left(\mathbf{X}_s, \boldsymbol{y}, s \right) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}s \\
&= q \frac{\partial}{\partial x_i} \frac{\partial \varphi}{\partial x_j} \int_\Omega \breve{\sigma}_{ik} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}t
\end{aligned}$$

Integrating by parts one has

$$\begin{aligned}
\int_\Omega \langle q \nabla \cdot \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t, \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle \, \mathrm{d}\boldsymbol{x} &= \int_\Omega q \frac{\partial}{\partial x_i} \frac{\partial \varphi}{\partial x_j} \int_\Omega \breve{\sigma}_{ik} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{x} \\
&= -\int_\Omega \varphi \frac{\partial}{\partial x_j} \left[q \frac{\partial}{\partial x_i} \int_\Omega \breve{\sigma}_{ik} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}t \right] \mathrm{d}\boldsymbol{x} \\
&= -\int_\Omega \varphi \nabla \cdot [q \boldsymbol{\sigma}_t^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma}_t \right)] \, \mathrm{d}\boldsymbol{x}.
\end{aligned}$$

■

2.9.2 Transport of the Jacobian

Consider the usual change of reference from a moving volume to a reference one, that is

$$\mathrm{d} \int_{V_t} \mathrm{d}\boldsymbol{x} = \mathrm{d} \int_{V_0} J \left(\mathbf{X}_t, t \right) \, \mathrm{d}\boldsymbol{y} \tag{2.167}$$

where J is the Jacobian determinant of the change of variable. However, employing the SRTT on the left hand side

$$\begin{aligned}
\mathrm{d} \int_{V_0} J \left(\mathbf{X}_t, t \right) \, \mathrm{d}\boldsymbol{y} &= \int_{V_t} \nabla \cdot [(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma} \right)) \, \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\mathbf{B}_t] \, \mathrm{d}\boldsymbol{x} \\
&= \int_{V_0} J \nabla \cdot [(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma} \right)) \, \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\mathbf{B}_t] \, \mathrm{d}\boldsymbol{y}
\end{aligned}$$

and thus thanks to the arbitrariness of the control volume V_t one has

$$\mathbb{D}_t J - J \nabla \cdot [(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}^{\mathrm{T}} \left(\nabla \cdot \boldsymbol{\sigma} \right)) \, \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\mathbf{B}_t] = 0. \tag{2.168}$$

2.9.3 Proof of distributivity of the stochastic transport operator

Starting from the application of the operator $\mathbb{D}_t$ to the product fg one has

$$\mathbb{D}_t(fg) = \mathrm{d}_t(fg) + [(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}_t^{\mathrm{T}}(\nabla \cdot \boldsymbol{\sigma}_t))\,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t] \cdot \nabla(fg) - \frac{1}{2}\nabla \cdot [\boldsymbol{a}\nabla(fg)]\,\mathrm{d}t \tag{2.169}$$

that can be expanded trivially as

$$\begin{aligned}\mathbb{D}_t(fg) =& g\left[\mathrm{d}_t f + [(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}_t^{\mathrm{T}}(\nabla \cdot \boldsymbol{\sigma}_t))\,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t] \cdot \nabla f - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla f)\,\mathrm{d}t\right] \\ &+ f\left[\mathrm{d}_t g + [(\boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}_t^{\mathrm{T}}(\nabla \cdot \boldsymbol{\sigma}_t))\,\mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t] \cdot \nabla g - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla g)\,\mathrm{d}t\right] \\ &- (\nabla f)^{\mathrm{T}}\,\boldsymbol{a}(\nabla g)\,\mathrm{d}t + \mathrm{d}\langle f, g\rangle_t\end{aligned}$$

where the last term stems from the application of Itô's integration by parts formula Kunita, 1997. This last term can be computed as

$$\begin{aligned}\mathrm{d}\langle f, g\rangle_t =& \langle \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla f, \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla g\rangle && (2.170)\\ &- \langle \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla f, \boldsymbol{G}_\sigma \cdot \mathrm{d}\mathbf{B}_t\rangle && (2.171)\\ &- \langle \boldsymbol{F}_\sigma \cdot \mathrm{d}\mathbf{B}_t, \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla g\rangle && (2.172)\\ &+ \langle \boldsymbol{F}_\sigma \cdot \mathrm{d}\mathbf{B}_t, \boldsymbol{G}_\sigma \cdot \mathrm{d}\mathbf{B}_t\rangle && (2.173)\\ =& (\nabla f)^{\mathrm{T}}\,\boldsymbol{a}(\nabla g)\,\mathrm{d}t && (2.174)\\ &- \boldsymbol{\sigma}_t \boldsymbol{G}_\sigma \cdot \nabla f\,\mathrm{d}t && (2.175)\\ &- \boldsymbol{\sigma}_t \boldsymbol{F}_\sigma \cdot \nabla g\,\mathrm{d}t && (2.176)\\ &+ \boldsymbol{F}_\sigma \cdot \boldsymbol{G}_\sigma\,\mathrm{d}t, && (2.177)\end{aligned}$$

that finally proves Equation (2.32).

Proof of equivalence between 2.170 and 2.174.

$$\begin{aligned}\langle \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla f, \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla g\rangle &\overset{\mathrm{def}}{=} \mathrm{d}\left\langle \int_0^{\cdot}\int_\Omega \frac{\partial f}{\partial x_i}\breve{\sigma}_{ik}(\mathbf{X}_s, \boldsymbol{y}, s)\,\mathrm{dB}_s^k(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \int_0^{\cdot}\int_\Omega \frac{\partial g}{\partial x_j}\breve{\sigma}_{j\ell}(\mathbf{X}_s, \boldsymbol{z}, s)\,\mathrm{dB}_s^\ell(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}\right\rangle_t \\ &= \mathrm{d}\int_0^t \int_{\Omega\times\Omega} \frac{\partial f}{\partial x_i}\breve{\sigma}_{ik}(\mathbf{X}_s, \boldsymbol{y}, s)\frac{\partial \theta}{\partial x_j}\breve{\sigma}_{j\ell}(\mathbf{X}_s, \boldsymbol{z}, s)\underbrace{\mathrm{d}\left\langle \mathrm{B}^k(\boldsymbol{y}), \mathrm{B}^\ell(\boldsymbol{z})\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s}\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \\ &= \mathrm{d}\int_0^t \int_\Omega \frac{\partial f}{\partial x_i}\breve{\sigma}_{ik}(\mathbf{X}_s, \boldsymbol{y}, s)\frac{\partial \theta}{\partial x_j}\breve{\sigma}_{jk}(\mathbf{X}_s, \boldsymbol{y}, s)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}s \\ &= (\nabla f)^{\mathrm{T}}\,\boldsymbol{a}(\nabla g)\,\mathrm{d}t\end{aligned}$$

■

Proof of equivalence between 2.171 and 2.175, and between 2.172 and 2.176.

$$\begin{aligned}\langle \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla f, \boldsymbol{G}_\sigma \cdot \mathrm{d}\mathbf{B}_t \rangle &\overset{\text{def}}{=} \mathrm{d}\left\langle \int_0^{\cdot} \int_\Omega \frac{\partial f}{\partial x_i} \breve{\sigma}_{ik}\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \mathrm{dB}_s^k\left(\boldsymbol{y}\right) \mathrm{d}\boldsymbol{y}, \int_0^{\cdot} \int_\Omega \breve{G}_\sigma^\ell\left(\mathbf{X}_s, \boldsymbol{z}, s\right) \mathrm{dB}_s^\ell\left(\boldsymbol{z}\right) \mathrm{d}\boldsymbol{z} \right\rangle_t \\ &= \mathrm{d}\int_0^t \int_{\Omega\times\Omega} \frac{\partial f}{\partial x_i} \breve{\sigma}_{ik}\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \breve{G}_\sigma^\ell\left(\mathbf{X}_s, \boldsymbol{z}, s\right) \underbrace{\mathrm{d}\left\langle \mathrm{B}^k\left(\boldsymbol{y}\right), \mathrm{B}^\ell\left(\boldsymbol{z}\right)\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s} \mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \\ &= \mathrm{d}\int_0^t \int_\Omega \frac{\partial f}{\partial x_i} \breve{\sigma}_{ik}\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \breve{G}_\sigma^k\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \mathrm{d}\boldsymbol{y}\,\mathrm{d}s \\ &= \boldsymbol{\sigma}_t \boldsymbol{G}_\sigma \cdot \nabla f \,\mathrm{d}t\end{aligned}$$

■

Proof of equivalence between 2.173 and 2.177.

$$\begin{aligned}\langle \boldsymbol{F}_\sigma \cdot \mathrm{d}\mathbf{B}_t, \boldsymbol{G}_\sigma \cdot \mathrm{d}\mathbf{B}_t \rangle &\overset{\text{def}}{=} \mathrm{d}\left\langle \int_0^{\cdot} \int_\Omega \breve{F}_\sigma^k\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \mathrm{dB}_s^k\left(\boldsymbol{y}\right) \mathrm{d}\boldsymbol{y}, \int_0^{\cdot} \int_\Omega \breve{G}_\sigma^\ell\left(\mathbf{X}_s, \boldsymbol{z}, s\right) \mathrm{dB}_s^\ell\left(\boldsymbol{z}\right) \mathrm{d}\boldsymbol{z} \right\rangle \\ &= \mathrm{d}\int_0^t \int_{\Omega\times\Omega} \breve{F}_\sigma^k\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \breve{G}_\sigma^\ell\left(\mathbf{X}_s, \boldsymbol{z}, s\right) \underbrace{\mathrm{d}\left\langle \mathrm{B}^k\left(\boldsymbol{y}\right), \mathrm{B}^\ell\left(\boldsymbol{z}\right)\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s} \mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} \\ &= \mathrm{d}\int_0^t \int_\Omega \breve{F}_\sigma^k\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \breve{G}_\sigma^k\left(\mathbf{X}_s, \boldsymbol{y}, s\right) \mathrm{d}\boldsymbol{y}\,\mathrm{d}s \\ &= \boldsymbol{F}_\sigma \cdot \boldsymbol{G}_\sigma \,\mathrm{d}t\end{aligned}$$

■

2.9.4 Stochastic transport theorem in advection form for an extensive tracer

Starting from Equation (2.47), that is

$$\begin{aligned}\mathrm{d}_t\left(\rho\theta\right) + \nabla\cdot\left[\rho\theta\left(\boldsymbol{v}-\boldsymbol{v}^s\right)\mathrm{d}t + \rho\theta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right] - \frac{1}{2}\nabla\cdot\left[\boldsymbol{a}\nabla\left(\rho\theta\right)\right]\mathrm{d}t + \nabla\cdot\left(\rho\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma\right)\mathrm{d}t \\ = \rho\Theta_t\,\mathrm{d}t + \rho\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t,\end{aligned}$$

one has to develop every term in order to split the two variables ρ and θ. The divergence of the flux can be split as

$$\begin{aligned}\nabla\cdot\left[\rho\theta\left(\boldsymbol{v}-\boldsymbol{v}^s\right)\mathrm{d}t + \rho\theta\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right] &= \theta\left(\nabla\cdot\left[\rho\left(\boldsymbol{v}-\boldsymbol{v}^s\right)\mathrm{d}t + \rho\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]\right) \\ &\quad + \rho\left(\left[\left(\boldsymbol{v}-\boldsymbol{v}^s\right)\mathrm{d}t + \boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t\right]\cdot\nabla\theta\right), \qquad (2.178)\end{aligned}$$

while the stochastic diffusion is unpacked as

$$\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\rho\theta)\,\mathrm{d}t=\rho\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}t+(\nabla\rho)^{\mathrm{T}}\,\boldsymbol{a}\,(\nabla\theta)+\theta\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\rho)\,\mathrm{d}t. \tag{2.179}$$

The temporal evolution of $\rho\theta$ is split considering Itô integration by parts formula Kunita, 1997 as

$$\mathrm{d}_t\left(\rho\theta\right)=\theta\mathrm{d}_t\rho+\rho\mathrm{d}_t\theta+\mathrm{d}\langle\rho,\theta\rangle_t. \tag{2.180}$$

The original Equation (2.47) can thus be rewritten as:

$$\begin{aligned}\mathrm{d}_t\int_\Omega\rho\theta\,\mathrm{d}\boldsymbol{x}=&\int_\Omega\theta\left[\mathrm{d}_t\rho+\nabla\cdot[\rho\left(\boldsymbol{v}-\boldsymbol{v}^s\right)\,\mathrm{d}t+\rho\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t]-\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\rho)\,\mathrm{d}t\right]\,\mathrm{d}\boldsymbol{x}\\&+\int_\Omega\rho\left[\mathrm{d}_t\theta+[(\boldsymbol{v}-\boldsymbol{v}^s)\,\mathrm{d}t+\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t]\cdot\nabla\theta-\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}t\right]\,\mathrm{d}\boldsymbol{x}\\&-\int_\Omega\rho\left[\Theta_t\,\mathrm{d}t+\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t\right]\,\mathrm{d}\boldsymbol{x}\\&+\int_\Omega\left[\mathrm{d}\langle\rho,\theta\rangle_t-(\nabla\rho)^{\mathrm{T}}\,\boldsymbol{a}\,(\nabla\theta)\,\mathrm{d}t+\nabla\cdot(\rho\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\,\mathrm{d}t\right]\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Thanks to the (assumed) conservation of the density ρ, the first integral vanishes as the content of the square bracket is identically zero. The quadratic covariation term can be shown to proved four terms, as

$$\begin{aligned}\mathrm{d}\langle\rho,\theta\rangle_t=&\langle-\nabla\cdot(\rho\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t),-\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\theta\rangle\\&+\langle\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t,-\nabla\cdot(\rho\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t)\rangle\\=&\langle\nabla\cdot(\rho\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t),\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\theta\rangle &(2.181)\\&-\langle\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t,\nabla\cdot(\rho\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t)\rangle &(2.182)\\=&\rho\left[\boldsymbol{\sigma}_t^{\mathrm{T}}\,(\nabla\cdot\boldsymbol{\sigma}_t)\cdot\nabla\theta\right]\,\mathrm{d}t+(\nabla\rho)^{\mathrm{T}}\,\boldsymbol{a}\,(\nabla\theta)\,\mathrm{d}t &(2.183)\\&-\left[\nabla\cdot(\rho\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)-\rho\mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma\right)\right]\mathrm{d}t, &(2.184)\end{aligned}$$

so that after cancellation one finds

$$\int_\Omega\rho\left[\mathrm{d}_t\theta+[\boldsymbol{v}^\star\,\mathrm{d}t+\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t]\cdot\nabla\theta-\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}t+\mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma\right)\,\mathrm{d}t-\Theta_t\,\mathrm{d}t-\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t\right]\,\mathrm{d}\boldsymbol{x}=0$$

with $\boldsymbol{v}^\star = \boldsymbol{v} - \boldsymbol{v}^s + \boldsymbol{\sigma}_t^{\mathrm{T}}\left(\nabla\cdot\boldsymbol{\sigma}_t\right)$. Using the definition of the stochastic transport operator one can rearrange the terms as

$$\int_\Omega \rho\left[\mathbb{D}_t\theta + \mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma\right)\,\mathrm{d}t - \Theta_t\,\mathrm{d}t - \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t\right]\mathrm{d}\boldsymbol{x} = 0,$$

which provides (2.49) straightforwardly.

Proof of equivalence between 2.181 and 2.183.

$$\begin{aligned}
\left\langle\nabla\cdot\left(\rho\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\right),\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\cdot\nabla\theta\right\rangle &\overset{\text{def}}{=} \mathrm{d}\left\langle\int_0^\cdot\int_\Omega\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{dB}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y},\int_0^\cdot\int_\Omega\frac{\partial\theta}{\partial x_j}\breve{\sigma}_{j\ell}\left(\mathbf{X}_s,\boldsymbol{z},s\right)\,\mathrm{dB}_s^\ell\left(\boldsymbol{z}\right)\,\mathrm{d}\boldsymbol{z}\right\rangle_t\\
&= \mathrm{d}\int_0^t\int_{\Omega\times\Omega}\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial\theta}{\partial x_j}\breve{\sigma}_{j\ell}\left(\mathbf{X}_s,\boldsymbol{z},s\right)\underbrace{\mathrm{d}\left\langle\mathrm{B}^k\left(\boldsymbol{y}\right),\mathrm{B}^\ell\left(\boldsymbol{z}\right)\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s}\,\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z}\\
&= \mathrm{d}\int_0^t\int_\Omega\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial\theta}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}s\\
&= \int_\Omega\left[\frac{\partial\rho}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)+\rho\frac{\partial}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\right]\frac{\partial\theta}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t\\
&= \int_\Omega\frac{\partial\rho}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\frac{\partial\theta}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t\\
&\quad+\int_\Omega\rho\frac{\partial}{\partial x_i}\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\frac{\partial\theta}{\partial x_j}\breve{\sigma}_{jk}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t\\
&= \rho\left[\boldsymbol{\sigma}_t^{\mathrm{T}}\left(\nabla\cdot\boldsymbol{\sigma}_t\right)\cdot\nabla\theta\right]\,\mathrm{d}t+\left(\nabla\rho\right)^{\mathrm{T}}\boldsymbol{a}\left(\nabla\theta\right)\,\mathrm{d}t
\end{aligned}$$

■

Proof of equivalence between 2.182 and 2.184.

$$\begin{aligned}
\left\langle\boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\mathbf{B}_t,\nabla\cdot\left(\rho\boldsymbol{\sigma}_t\,\mathrm{d}\mathbf{B}_t\right)\right\rangle &\overset{\text{def}}{=} \mathrm{d}\left\langle\int_0^\cdot\int_\Omega\breve{\Theta}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{dB}_s^k\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y},\int_0^\cdot\int_\Omega\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{i\ell}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{dB}_s^\ell\left(\boldsymbol{y}\right)\,\mathrm{d}\boldsymbol{y}\right\rangle_t\\
&= \mathrm{d}\int_0^t\int_{\Omega\times\Omega}\breve{\Theta}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{i\ell}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\underbrace{\mathrm{d}\left\langle\mathrm{B}^k\left(\boldsymbol{y}\right),\mathrm{B}^\ell\left(\boldsymbol{z}\right)\right\rangle_s}_{\delta_{k\ell}\delta(\boldsymbol{y}-\boldsymbol{z})\,\mathrm{d}s}\,\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z}\\
&= \mathrm{d}\int_0^t\int_\Omega\breve{\Theta}_\sigma^k\left(\mathbf{X}_s,\boldsymbol{y},s\right)\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{ik}\left(\mathbf{X}_s,\boldsymbol{y},s\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}s\\
&= \int_\Omega\breve{\Theta}_\sigma^k\left(\mathbf{X}_t,\boldsymbol{y},t\right)\frac{\partial}{\partial x_i}\rho\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}t\\
&= \int_\Omega\frac{\partial}{\partial x_i}\left(\rho\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\breve{\Theta}_\sigma^k\left(\mathbf{X}_t,\boldsymbol{y},t\right)\right)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}t\\
&\quad-\int_\Omega\rho\breve{\sigma}_{ik}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\frac{\partial\breve{\Theta}_\sigma^k}{\partial x_i}\left(\mathbf{X}_t,\boldsymbol{y},t\right)\,\mathrm{d}\boldsymbol{y}\mathrm{d}t\\
&= \left[\nabla\cdot\left(\rho\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma\right)-\rho\mathrm{Tr}\left(\boldsymbol{\sigma}^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma\right)\right]\mathrm{d}t
\end{aligned}$$

■

2.9.5 Quadratic variation of pressure gradient

$$\begin{aligned}
\langle \nabla \mathrm{d}p_t^\sigma, \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle &\overset{\text{def}}{=} \mathrm{d} \left\langle \int_0^\cdot \frac{\partial}{\partial x_d} \int_\Omega \vartheta_k \left(\mathbf{X}_s, \boldsymbol{y}, t \right) \mathrm{d}\mathrm{B}_s^k \left(\boldsymbol{y} \right) \mathrm{d}\boldsymbol{y}, \int_0^\cdot \int_\Omega \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{j\ell} \left(\mathbf{X}_s, \boldsymbol{z}, s \right) \mathrm{d}\mathrm{B}_s^\ell \left(\boldsymbol{z} \right) \mathrm{d}\boldsymbol{z} \right\rangle_t \\
&= \mathrm{d} \int_0^t \int_{\Omega \times \Omega} \frac{\partial \vartheta_k \left(\mathbf{X}_s, \boldsymbol{y}, s \right)}{\partial x_d} \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{j\ell} \left(\mathbf{X}_s, \boldsymbol{z}, s \right) \underbrace{\mathrm{d} \left\langle \mathrm{B}^k \left(\boldsymbol{y} \right), \mathrm{B}^\ell \left(\boldsymbol{z} \right) \right\rangle_s}_{\delta_{k\ell} \delta(\boldsymbol{y} - \boldsymbol{z}) \, \mathrm{d}s} \mathrm{d}\boldsymbol{y} \mathrm{d}\boldsymbol{z} \\
&= \mathrm{d} \int_0^t \int_\Omega \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{jk} \left(\mathbf{X}_s, \boldsymbol{y}, s \right) \frac{\partial \vartheta_k \left(\mathbf{X}_s, \boldsymbol{y}, s \right)}{\partial x_d} \mathrm{d}\boldsymbol{y} \mathrm{d}s \\
&= \int_\Omega \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \frac{\partial \vartheta_k \left(\mathbf{X}_s, \boldsymbol{y}, t \right)}{\partial x_d} \mathrm{d}\boldsymbol{y} \mathrm{d}t
\end{aligned}$$

Integrating by parts

$$\begin{aligned}
- \int_\Omega \langle \nabla \mathrm{d}p_t^\sigma, \boldsymbol{\sigma}_t \, \mathrm{d}\mathbf{B}_t \cdot \nabla \varphi \rangle \, \mathrm{d}\boldsymbol{x} &= - \int_\Omega \int_\Omega \frac{\partial \varphi}{\partial x_j} \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \frac{\partial \vartheta_k \left(\mathbf{X}_t, \boldsymbol{y}, t \right)}{\partial x_d} \mathrm{d}\boldsymbol{y} \mathrm{d}t \, \mathrm{d}\boldsymbol{x} \\
&= - \int_\Omega \frac{\partial \varphi}{\partial x_j} \int_\Omega \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \frac{\partial \vartheta_k \left(\mathbf{X}_t, \boldsymbol{y}, t \right)}{\partial x_d} \mathrm{d}\boldsymbol{y} \mathrm{d}t \, \mathrm{d}\boldsymbol{x} \\
&= \int_\Omega \varphi \frac{\partial}{\partial x_j} \int_\Omega \breve{\sigma}_{jk} \left(\mathbf{X}_t, \boldsymbol{y}, t \right) \frac{\partial \vartheta_k \left(\mathbf{X}_t, \boldsymbol{y}, t \right)}{\partial x_d} \mathrm{d}\boldsymbol{y} \mathrm{d}t \, \mathrm{d}\boldsymbol{x} \\
&= \int_\Omega \varphi \nabla \cdot \left[\left(\nabla \boldsymbol{\vartheta} \right) \boldsymbol{\sigma}_t \right] \mathrm{d}t \, \mathrm{d}\boldsymbol{x}
\end{aligned}$$

where in the last line symmetry of the kernel $\breve{\sigma}$ was used.

2.9.6 Stochastic non inertial acceleration

Applying Frenet-Serret formula (2.99) to the velocity Equation (2.7) one has

$$\begin{aligned}
\mathrm{d}_I \boldsymbol{v} = \mathrm{d}_I \boldsymbol{v}' &= \mathrm{d}_I \left(v_1' \boldsymbol{e}_1' + v_2' \boldsymbol{e}_2' + v_3' \boldsymbol{e}_3' \right) \\
&= \left(\mathrm{d}_I v_i' \right) \boldsymbol{e}_i' + v_i' \mathrm{d}_I \boldsymbol{e}_i' + \frac{1}{2} \left\langle v_i', e_i' \right\rangle \\
&= \left(\mathrm{d}_R v_i' \right) \boldsymbol{e}_i' + \boldsymbol{\Omega} \, \mathrm{d}t \times \left(v_i' \boldsymbol{e}_i' \right) \\
&= \mathrm{d}_R \boldsymbol{v}' + \left(\boldsymbol{\Omega} \, \mathrm{d}t \times \boldsymbol{v}' \right).
\end{aligned}$$

Applying again the definition (2.7) to the momentum variation

$$\begin{aligned}
\mathrm{d}_I \int_{V_t} \rho \boldsymbol{V} \, \mathrm{d}\boldsymbol{x} &= \mathrm{d}_R \int_{V_t} \rho \boldsymbol{V}' \, \mathrm{d}\boldsymbol{x} + \boldsymbol{\Omega} \, \mathrm{d}t \times \left(\int_{V_t} \rho \boldsymbol{V}' \, \mathrm{d}\boldsymbol{x} \right) \\
\mathrm{d}_I \int_{V_t} \rho \left(\boldsymbol{v} + \boldsymbol{\sigma} \dot{\mathbf{B}}_t \right) \mathrm{d}\boldsymbol{x} &= \mathrm{d}_R \int_{V_t} \rho \left(\boldsymbol{v}' + \boldsymbol{\sigma}' \dot{\mathbf{B}}_t \right) \mathrm{d}\boldsymbol{x} + \boldsymbol{\Omega} \, \mathrm{d}t \times \int_{V_t} \rho \left(\boldsymbol{v}' + \boldsymbol{\sigma}' \dot{\mathbf{B}}_t \right) \mathrm{d}\boldsymbol{x}
\end{aligned}$$

and interpreting the derivatives in distributional senses one finds the relation

$$\int_T \varphi(s) \left[\mathrm{d}_I \int_{V_s} \rho \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} \right] \mathrm{d}s = \int_T \varphi(s) \left[\mathrm{d}_R \int_{V_s} \rho \boldsymbol{v}' \, \mathrm{d}\boldsymbol{x} + \boldsymbol{\Omega} \mathrm{d}t \times \int_{V_s} \rho \left(\boldsymbol{v}' \, \mathrm{d}s + \boldsymbol{\sigma}' \mathrm{d}\mathbf{B}_s \right) \mathrm{d}\boldsymbol{x} \right] \mathrm{d}s - \int_T \varphi'(s) \int_{V_s} \rho \boldsymbol{\sigma} \mathrm{d}\mathbf{B}_s \, \mathrm{d}\boldsymbol{x}.$$

that can be equated to the source of momentum variations as in sections 2.4.3 and 2.5.

Chapter 3

Noise modelling

Abstract

This chapter reviews the main techniques that has been used to construct the noise models employed in this study. The theoretical properties of each model are presented in a light-weight manner, as most of them are classical techniques well documented in the literature. Where the particular model was not suited to be applied to this study or it required further, non-classical modifications, those are exposed and explained.

The noise $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t$ is defined through the action of an integration kernel $\breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{y}, t)$ on a 3D functional Brownian motion $\mathrm{d}\mathbf{B}_t$. The integration kernel encodes the desired properties that must be satisfied by the small scale stochastic contribution. In particular, $\breve{\sigma}$ is required to encode the spatial correlation of the small scale turbulent motions, so the operator

$$\begin{aligned} \boldsymbol{\sigma}_t \colon \left[L^2(\Omega)\right]^d &\to \left[L^2(\Omega)\right]^d \\ \boldsymbol{g} &\mapsto \int_\Omega \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{y}, t)\, g^k(\boldsymbol{y})\, d\boldsymbol{y} \end{aligned} \tag{3.1}$$

is defined as a symmetric operator having a bounded kernel (Hilbert-Schmidt operator). The application of such a kernel to a cylindrical Wiener process $\mathbf{B}_t^k$ results in a centered Gaussian process with one-point, one-time covariance, denoted as $\boldsymbol{a}$, given by the diagonal components of the two-point, two-time covariance tensor per unit of time as introduced in Chapter 2, that is

$$a_{ij}(\boldsymbol{x}, t) = \int_\Omega \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{y}, t)\, \breve{\sigma}_{kj}(\boldsymbol{x}, \boldsymbol{y}, t)\, \mathrm{d}\boldsymbol{y}. \tag{3.2}$$

This variance tensor is symmetric. The choice of having a bounded integration kernel, thus being a Hilbert-Schmidt operator, allow us to define the noise and variance term as

the expansions

$$\boldsymbol{\sigma}_t(\boldsymbol{x})\,\mathrm{d}\mathbf{B}_t = \sum_{n\in\mathbb{N}} \sqrt{\lambda_n}\,\boldsymbol{\phi}_n(\boldsymbol{x})\,\mathrm{d}\beta_t^n, \quad \boldsymbol{a}(\boldsymbol{x},t) = \sum_{n\in\mathbb{N}} \lambda_n\,\boldsymbol{\phi}_n(\boldsymbol{x})\,\boldsymbol{\phi}_n^{\mathrm{T}}(\boldsymbol{x}), \tag{3.3}$$

where $\{\beta_t^n, n \in \mathbb{N}\}$ is a set of 1D standard (scalar) Brownian motions. This representation corresponds to the Karhunen-Loeve decomposition (Loeve, 1978). In the following, operatives approaches to define the noise are going to be presented. In particular, as the kernel $\breve{\sigma}$ is wanted to encode spatial characteristics of the flow, different techniques will be introduced from the general framework of approximation theory in order to build a set of basis functions and amplitudes $\{\boldsymbol{\phi}_n(\boldsymbol{x}), \lambda_n\}_{n\in\mathbb{N}}$ to produce a physically meaningful noise. Suppose that a real valued function $\boldsymbol{f}(\boldsymbol{x},\boldsymbol{t})$ taking values over a domain $\mathcal{D} = \Omega \times [0,T]$ has to be approximated. The classical starting point is to define an approximation $\boldsymbol{f}_K$ such that

$$\boldsymbol{f}_K(\boldsymbol{x},t) \simeq \sum_{k=1}^{K} a_k(t)\,\boldsymbol{\phi}_k(\boldsymbol{x}) \tag{3.4}$$

with the expectation that

$$\lim_{K\to\infty} |\boldsymbol{f} - \boldsymbol{f}_K| = 0. \tag{3.5}$$

On one hand, if the function $\boldsymbol{f}$ is unknown, the set of basis functions $\{\boldsymbol{\phi}_k(\boldsymbol{x})\}_{k=1}^N$ is specified up to a given degree N and the approximated solution $\boldsymbol{f}_N$ is used to replace the target function in the equations describing the process, generally resulting in a linear system in the unknowns a_k that, once solved, is used to define completely the approximation $\boldsymbol{f}_N$ of $\boldsymbol{f}$. Depending on choice of $\{\boldsymbol{\phi}_k(\boldsymbol{x})\}_{k=1}^N$ one can define different methods and applications: locally supported polynomials are used in Finite Element Analysis (Zienkiewicz et al., 2013), Legendre or Chebyshev polynomials are employed for Spectral Methods (Canuto et al., 1988), Radial Basis functions and Reproducing Kernel Hilbert spaces for Meshless methods (Fasshauer, 2007). On the other hand, when the function $\boldsymbol{f}$ is somehow known then the set of basis functions $\{\boldsymbol{\phi}_k(\boldsymbol{x})\}_{k=1}^N$ can be built taking into account the intrinsic nature of the function $\boldsymbol{f}$. This class of approaches includes *proper orthogonal decomposition* (POD), *dynamic mode decomposition* (DMD), *wavelet transform* (WLT) and is generally referred to as *feature extraction*, or *a posteriori* methods. The operative procedure chosen is thus the latter, assuming that a function $\boldsymbol{f}$ is known and it is representative of the flow. The choice of the function $\boldsymbol{f}$ is however non-trivial, as if it is asked to encode and represent all the processes that cannot be simulated then it must be a function of the state variables of interest (*e.g.* velocity and active tracers like temperature and salinity, in

the case of geophysical fluid mechanics) and must reflect the kind of processes that must be stochastically modelled.

3.1 Data filtering

If the aim is that of improving a fine resolution simulation to enhance its representation capabilities in a statistical sense, then the target function $\boldsymbol{f}$ might come from the current state of the simulation. In this case, temporal averaging is not always convenient, as the time average is only capturing the past states as

$$\overline{\boldsymbol{u}}^{T}(\boldsymbol{x},t)=\frac{1}{T}\int_{t-T}^{t}\boldsymbol{u}(\boldsymbol{x},\tau)\,\mathrm{d}\tau. \tag{3.6}$$

A spatial filtering is thus applied instead, defined as

$$\overline{\boldsymbol{u}}^{\nu}(\boldsymbol{x},t)=\int_{\nu(\boldsymbol{x})}\boldsymbol{u}(\boldsymbol{\xi},t)\,\varpi(\boldsymbol{\xi},t)\,\mathrm{d}\boldsymbol{\xi}, \tag{3.7}$$

where $\nu(\boldsymbol{x})$ is a neighbourhood of the point $\boldsymbol{x}$ and $\varpi(\boldsymbol{\xi},t)$ is a weighting function. The weighting function can be chosen in such a way that the resulting integral corresponds to a classical average in space or to a Gaussian filtering.

3.2 On the vertical structure of the noise

Before embarking in the description of each noise model, it is worth remarking one fundamental aspect common to all the models. Equations (2.117a,b) prescribe incompressibility of both the large scale velocity and noise field. The noise to be inserted into the model must thus be incompressible. This can be done straightforwardly by constructing the noise in the horizontal direction only and then defining its vertical component by integrating from bottom to the free surface its horizontal divergence, so that

$$\sigma_t\mathrm{d}\mathrm{B}_t^z(\boldsymbol{x},z)=-\int_{\eta_b}^{z}\nabla_{\mathrm{H}}\cdot\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{\mathrm{H}}(\boldsymbol{x},\zeta)\,\mathrm{d}\zeta. \tag{3.8}$$

For this reason, in the following the velocity fields will be denoted by $\boldsymbol{u}$, indicating the horizontal velocity in accordance to the splitting $\boldsymbol{v}=(\boldsymbol{u},w)$. The focus on this chapter is on the procedure to create the different noise models, it is thus implied in the description

of the noise ansatz that the vertical component is constructed starting from the horizontal one in accordance with Equation (3.8).

3.3 Offline data filtering

If the aim is that of improving a coarse resolution simulation with respect to a higher resolution, the target function $\boldsymbol{f}$ can be chosen as the residual between a coarse resolution simulation and a fine resolution one. The high resolution data used to force the low resolution stochastic model need to be filtered before being used, in order to avoid the injection of energy scales that can jeopardise the stability of the simulation. The low resolution velocity fluctuations are obtained through spatial filtering of high resolution temporal fluctuations. First, a time average of type (3.6) is applied on the high resolution fields as

$$\overline{\boldsymbol{u}}^{T}_{\text{HR}}\left(\boldsymbol{x},t\right)=\frac{1}{T}\int_{t-T}^{t}\boldsymbol{u}_{\text{HR}}\left(\boldsymbol{x},s\right)\,\mathrm{d}s, \tag{3.9}$$

so to obtain with Reynolds decomposition the high resolution fluctuations:

$$\boldsymbol{u}'_{\text{HR}}\left(\boldsymbol{x},t\right)=\boldsymbol{u}_{\text{HR}}\left(\boldsymbol{x},t\right)-\overline{\boldsymbol{u}}^{T}_{\text{HR}}\left(\boldsymbol{x},t\right). \tag{3.10}$$

The corresponding low resolution fluctuations are obtained through a band-pass filter $\mathcal{B}_{1,2}$ that is defined through a combination of two a Gaussian filters of type (3.7) (with ϖ Gaussian weights) and a successive down-sampling to the low resolution grid as $\mathcal{B}^{\text{LR}}_{1,2}=(\mathcal{G}^{\text{LR}}_{1}-\mathcal{G}^{\text{LR}}_{2})$, so that the fluctuations read

$$\boldsymbol{u}'_{\text{LR}}=\left(\mathcal{G}^{\text{LR}}_{1}-\mathcal{G}^{\text{LR}}_{2}\right)\boldsymbol{u}'_{\text{HR}}. \tag{3.11}$$

This filtered and coarse grained field has a smaller amount of energy compared to the original. In the case the energy level of the initial field has to be retained, a scaling can be applied as

$$\boldsymbol{u}'_{\text{LR}}=\frac{\left\|\boldsymbol{u}'_{\text{HR}}\right\|_{2}}{\left\|\mathcal{B}^{\text{LR}}_{1,2}\boldsymbol{u}'_{\text{HR}}\right\|_{2}}\mathcal{B}^{\text{LR}}_{1,2}\boldsymbol{u}'_{\text{HR}}, \tag{3.12}$$

The result of this procedure sees the velocity fluctuations have the same spatial structure as before but enhanced level of energy.

3.4 Data decomposition

The general formalism that will be adopted in the following sections is here introduced. Let the time domain $\mathbb{R}^+$ be sampled with a uniform time discretization $\{t_k = (k-1)\,\Delta t\}_{k=1}^{n_t}$. The spatial domain $\Omega \subset \mathbb{R}^d$ is sampled on a Cartesian grid $\boldsymbol{x}_i \in \mathbb{R}^{n_1 \times \ldots \times n_d}$, with $i \in (1, \ldots, n_s)$ a matrix linear index, such that $n_s = n_c n_1 \cdots n_d$ and n_c the number of components of the system ($n_c = 1$ for a scalar quantity, $n_c = 2$ for a horizontal velocity or $n_c = 3$ for a three dimensional velocity). Observations of a random process $\{\boldsymbol{d}(\boldsymbol{x}, t)\,,\ \boldsymbol{x} \in \Omega, t \in \mathbb{R}^+\}$ over the Cartesian grid $\boldsymbol{x}_i$ at times t_k are collected into the matrix $\boldsymbol{D} \in \mathbb{R}^{n_s \times n_t}$ in such a way that the matrix reads

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{d}_1(x_1) & \boldsymbol{d}_2(x_1) & \ldots & \boldsymbol{d}_{n_t}(x_1) \\ \vdots & \vdots & & \vdots \\ \boldsymbol{d}_1(x_{n_s}) & \boldsymbol{d}_2(x_{n_s}) & \ldots & \boldsymbol{d}_{n_t}(x_{n_s}) \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \boldsymbol{d}_1 & \boldsymbol{d}_2 & \ldots & \boldsymbol{d}_{n_t} \\ | & | & & | \end{bmatrix}. \tag{3.13}$$

Any decomposition aims at describing the process $\boldsymbol{d}$ as in separated variable form, that means as the result of a spatial contribution $\boldsymbol{\phi}$, a temporal contribution $\boldsymbol{\psi}$ and an amplitude σ, that is

$$\boldsymbol{D}(\boldsymbol{x}, t) = \sum_{r=1}^{\mathrm{rk}(\boldsymbol{D})} \sigma_r \boldsymbol{\phi}_r(\boldsymbol{x})\, \boldsymbol{\psi}_r(t)\,, \tag{3.14}$$

that in matrix notation becomes

$$\boldsymbol{D} = \sum_{r=1}^{\mathrm{rk}(\boldsymbol{D})} \sigma_r \boldsymbol{\phi}_r \boldsymbol{\psi}_r^{\mathrm{T}} = \boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Psi}^{\mathrm{T}}, \tag{3.15}$$

where $\boldsymbol{\Sigma} = \mathrm{diag}\left(\sigma_1, \ldots, \sigma_{\mathrm{rk}(\boldsymbol{D})}\right)$ is the diagonal matrix containing the energy contribution of each mode, while $\boldsymbol{\Phi} = \left[\boldsymbol{\phi}_1(\boldsymbol{x}), \ldots, \boldsymbol{\phi}_{\mathrm{rk}(\boldsymbol{D})}(\boldsymbol{x})\right] \in \mathbb{R}^{n_s \times \mathrm{rk}(\boldsymbol{D})}$ and $\boldsymbol{\Psi} = \left[\boldsymbol{\psi}_1(t), \ldots, \boldsymbol{\psi}_{\mathrm{rk}(\boldsymbol{D})}(t)\right] \in \mathbb{R}^{n_t \times \mathrm{rk}(\boldsymbol{D})}$. Equation (3.15) shows that the operation that is performed is a projection of the dataset onto the spatial basis $\boldsymbol{\Phi}$ and a temporal basis $\boldsymbol{\Psi}$, the former for the columns of $\boldsymbol{D}$, the latter for its rows. These temporal structures $\boldsymbol{\Psi}$ are prescribed, so to focus on the time projection of the dataset and define $\boldsymbol{\Phi}$ as

$$\boldsymbol{\Phi} = \left(\boldsymbol{D}\left(\boldsymbol{\Psi}^{\mathrm{T}}\right)^{-1}\right)\boldsymbol{\Sigma}^{-1}. \tag{3.16}$$

The energy content can be computed as

$$\|\boldsymbol{D}\|_{\mathrm{F}}^2 = \mathrm{Tr}\left(\boldsymbol{K}\right) = \mathrm{Tr}\left(\boldsymbol{C}\right) = \sum_{r=1}^{\mathrm{rk}(\boldsymbol{D})} \lambda_r \tag{3.17}$$

where $\boldsymbol{K} = \boldsymbol{D}^\dagger \boldsymbol{D} \in \mathbb{R}^{n_t \times n_t}$ and $\boldsymbol{C} = \boldsymbol{D}\boldsymbol{D}^\dagger \in \mathbb{R}^{n_s \times n_s}$ are finite dimensional estimators of the two point temporal and spatial correlation tensors, with $A^\dagger$ being the Hermitian adjoint of A. In the previous equation, $\|\boldsymbol{D}\|_{\mathrm{F}}^2$ is the Frobenius norm of D, while λ_r is the $r-$th eigenvalue of $\boldsymbol{K}$ and $\boldsymbol{C}$. Inserting equation (3.15) in the previous, one obtains a relation between the energy of the data and the spatial and temporal structures, that is

$$\|\boldsymbol{D}\|_{\mathrm{F}}^2 = \mathrm{Tr}\left(\boldsymbol{\Psi}\boldsymbol{\Sigma}\boldsymbol{\Phi}^\dagger\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Psi}^\dagger\right) = \mathrm{Tr}\left(\boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Psi}^\dagger\boldsymbol{\Psi}\boldsymbol{\Sigma}\boldsymbol{\Phi}^\dagger\right) = \sum_{r=1}^{\mathrm{rk}(\boldsymbol{D})} \lambda_r \tag{3.18}$$

that states that the matrix $\boldsymbol{\Sigma}$ provides an estimate of the energy of the data if and only if the temporal and spatial structures are both orthogonal, that is $\boldsymbol{\Phi}^\dagger\boldsymbol{\Phi} = \boldsymbol{\Psi}^\dagger\boldsymbol{\Psi} = I$. To enforce this property the spatial and temporal structures must have unitary energy. On the assumed Cartesian grid and uniform temporal sampling the average energy of the the spatial and temporal structures can be computed with discretized version of classical norms as

$$\|\boldsymbol{\phi}_r\|_2^2 = \int_\Omega \boldsymbol{\phi}_r\left(\boldsymbol{x}\right)\overline{\boldsymbol{\phi}}_r\left(\boldsymbol{x}\right)\,\mathrm{d}\boldsymbol{x}, \tag{3.19}$$

$$\|\boldsymbol{\psi}_r\|_2^2 = \int_T \boldsymbol{\psi}_r\left(t\right)\overline{\boldsymbol{\psi}_r}\left(t\right)\,\mathrm{d}t, \tag{3.20}$$

and thus the spatial and temporal structures should be scaled by respectively $1/\sqrt{\Omega}$ and $1/\sqrt{T}$ to ensure $\|\boldsymbol{\phi}_r\|_2 = \|\boldsymbol{\psi}_r\|_2 = 1$ in an domain-averaged sense. It follows that amplitudes σ_r must be scaled by $\sqrt{\Omega T}$ in order to provide a decomposition that is independent from the spatio-temporal domain of work. From Equations (3.14) and (3.15) it is clear that taking the time average of the dataset corresponds to define one of the temporal structures as $\overline{\boldsymbol{\psi}}^t = 1/\sqrt{n_t} \in \mathbb{R}^{n_t \times 1}$ so that one can write

$$\boldsymbol{D}\left(\boldsymbol{x},t\right) = \frac{1}{n_t}\boldsymbol{D}\left(\boldsymbol{x},t\right) + \breve{\boldsymbol{D}}\left(\boldsymbol{x},t\right) = \overline{\sigma}^t\overline{\boldsymbol{\phi}}^t\left(\boldsymbol{x}\right)\overline{\boldsymbol{\psi}}^t + \sum_{r=1}^{\mathrm{rk}(\boldsymbol{D})-1} \breve{\sigma}_r\breve{\boldsymbol{\phi}}_r\left(\boldsymbol{x}\right)\breve{\boldsymbol{\psi}}_r\left(t\right), \tag{3.21}$$

with $\breve{\boldsymbol{D}}$ de-trended dataset (*i.e.* $\breve{\boldsymbol{D}} + \boldsymbol{D} - \dfrac{1}{n_t}\boldsymbol{D}$) and $\breve{\boldsymbol{\phi}}_r$ and $\breve{\boldsymbol{\psi}}_r$ corresponding spatial and temporal structures. $\overline{\boldsymbol{\phi}}^t$ and $\overline{\sigma}^t$ represent the spatial scale and energy of the time average.

It is important to state that, even though in this dissertation a Cartesian grid and a uniform sampling are used, thus justifying the framework introduced, a generalization to non Cartesian grids or non uniform samplings is possible by adapting the discrete inner products to include metric terms, or pre-scaling the data with metric terms.

3.5 Proper orthogonal decomposition (POD)

The proper orthogonal decomposition is introduced here in the general context of approximation theory following Chatterjee (2000), Rivlin (1969) and Cordier et al. (2003). Obtaining a compact representation of data may be pursued with the multi-variate statistical method known as *proper orthogonal decomposition* (POD). The target of the procedure is to reduce the number of intercorrelated variables to a smaller set of uncorrelated variables while retaining as much of the variation in the initial variables, that is finding a representing subspace of fixed dimension which is optimal in the sense that the error in the projection onto this subspace is minimized. This serves the twofold cause of order reduction and feature extraction of the so called *coherent structures*. Let $\{\boldsymbol{d}(\boldsymbol{x},t)\,,\ \boldsymbol{x}\in\Omega, t\in\mathbb{R}^+\}$ be a set of observations of a random process over a spatial domain Ω. A coherent structure, as defined by Lumley (1970), is a deterministic function $\boldsymbol{\phi}$ which is best correlated, on average, with the realizations of $\boldsymbol{d}$. In other words, the functions $\boldsymbol{\phi}$ are those functions that possess the largest mean-square projection on the observations $\boldsymbol{d}$, that is $|\langle \boldsymbol{d}, \boldsymbol{\phi}\rangle|^2$. The interest on the functions $\boldsymbol{\phi}$ is in their spatial structures, so the amplitude of these functions should not be of impact on the choice, hence they are chosen to be normalised as $\|\boldsymbol{\phi}\|^2 = 1$ and the projection itself must be normalised by the norm of the the function. One can define a subspace S spanned by a set of coherent structures $\boldsymbol{\phi}_j$, with $j = 1,\ldots,n$ and thus defining the projection of $\boldsymbol{d}$ onto S as

$$P_S\boldsymbol{d} = \sum_{j=1}^{n} \frac{\langle \boldsymbol{d}, \boldsymbol{\phi}_j\rangle}{\|\boldsymbol{\phi}_j\|^2}\boldsymbol{\phi}_j \tag{3.22}$$

and thus the minimization of the mean square projection of $\boldsymbol{d}$ onto S can be stated as

$$\min_{\phi} \overline{\|\boldsymbol{d} - P_S\boldsymbol{d}\|^2}^{X}, \tag{3.23}$$

where the overbar means averaging in some sense. In particular, proper orthogonal decomposition is designed to minimize the number n of basis functions needed in equation

(3.22).

3.5.1 Mathematical formulation

In this section, the development of Holmes, L., et al. (1996), C. W. Rowley (2002) and C. Rowley (2005) is followed, describing the POD procedure in the context of general Hilbert spaces. Let $\mathcal{H}$ be ah Hilbert space with inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and induced norm $\|\cdot\|_{\mathcal{H}}$. The set of functions $\left\{\boldsymbol{\phi}_j(\boldsymbol{x}) \in \mathcal{H} :\ j = 1,\ldots,n\right\}$ is defined as the one that maximise the $X-$averaged projection of $\boldsymbol{d}$ onto $\boldsymbol{\phi}$, that is

$$\max_{\boldsymbol{\phi}\in\mathcal{H}} \frac{\overline{|\langle \boldsymbol{d},\boldsymbol{\phi}\rangle_{\mathcal{H}}|^2}^{X}}{\|\boldsymbol{\phi}\|_{\mathcal{H}}^2}, \tag{3.24}$$

subject to the constraint $\|\boldsymbol{\phi}\|^2 = 1$, to close the problem. A functional $\mathcal{J}[\boldsymbol{\phi}]$ can be defined as

$$\mathcal{J}[\boldsymbol{\phi}] = \overline{|\langle \boldsymbol{d},\boldsymbol{\phi}\rangle_{\mathcal{H}}|^2}^{X} - \lambda\left(\|\boldsymbol{\phi}\|_{\mathcal{H}}^2 - 1\right), \tag{3.25}$$

including the constraint through a Lagrange multiplier. A Gateaux derivative is performed to set to zero the infinitesimal variations $\boldsymbol{\phi} + \epsilon\boldsymbol{\psi} \in \mathcal{H}$, with $\epsilon \in \mathbb{R}$, that means

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{J}[\boldsymbol{\phi}+\epsilon\boldsymbol{\psi}] &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\left[\overline{\langle \boldsymbol{d},\boldsymbol{\phi}+\epsilon\boldsymbol{\psi}\rangle_{\mathcal{H}}\langle \boldsymbol{d},\boldsymbol{\phi}+\epsilon\boldsymbol{\psi}\rangle_{\mathcal{H}}}^{X} - \lambda\langle\boldsymbol{\phi}+\epsilon\boldsymbol{\psi},\boldsymbol{\phi}+\epsilon\boldsymbol{\psi}\rangle_{\mathcal{H}}\right]\Big|_{\epsilon=0} \\ &= 2\,\overline{\langle \boldsymbol{d},\boldsymbol{\phi}\rangle_{\mathcal{H}}\langle \boldsymbol{d},\boldsymbol{\psi}\rangle_{\mathcal{H}}}^{X} - 2\lambda\langle\boldsymbol{\phi},\boldsymbol{\psi}\rangle_{\mathcal{H}} = 0.\end{aligned}$$

Assuming commutation is possible between the averaging $\overline{\cdot}^{X}$ and the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ one has

$$\left\langle \overline{\langle \boldsymbol{d},\boldsymbol{\phi}\rangle_{\mathcal{H}}\boldsymbol{d}}^{X} - \lambda\boldsymbol{\phi},\boldsymbol{\psi}\right\rangle_{\mathcal{H}} = 0 \tag{3.26}$$

corresponding to the eigenproblem

$$\mathcal{R}\boldsymbol{\phi} = \lambda\boldsymbol{\phi} \tag{3.27}$$

where $\mathcal{R}\boldsymbol{\phi} = \overline{\langle \boldsymbol{d},\boldsymbol{\phi}\rangle_{\mathcal{H}}\boldsymbol{d}}^{X}$. Considering the case where the Hilbert space $\mathcal{H}$ is L^2, a natural case in fluid mechanics as it represents functions with finite kinetic energy, the linear operator $\mathcal{R}\boldsymbol{\phi}$ is

$$\begin{aligned}\mathcal{R}\boldsymbol{\phi} = \overline{\langle \boldsymbol{d},\boldsymbol{\phi}\rangle_{L^2}\boldsymbol{d}}^{X} &= \overline{\int_\Omega \boldsymbol{d}(\boldsymbol{\xi},t)\,\boldsymbol{\phi}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\,\boldsymbol{d}(\boldsymbol{x},t)}^{X} \\ &= \int_\Omega \overline{\boldsymbol{d}(\boldsymbol{\xi},t)\,\boldsymbol{d}(\boldsymbol{x},t)}^{X}\boldsymbol{\phi}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}.\end{aligned}$$

that is the $X-$averaged two points autocorrelation function of $\boldsymbol{d}$. The functions are orthogonal in the sense that

$$\int_{\Omega} \boldsymbol{\phi}_j(\boldsymbol{x})\, \boldsymbol{\phi}_k(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \delta_{jk}, \tag{3.28}$$

It is crucial to notice that the temporally averaged quantity corresponds to the two points auto-correlation function of the signal. An orthogonal transformation is performed to project the data onto the subspace generated by the eigenvectors of the sample covariance matrix. This gives the optimal linear manifold approximating the data, in the sense that it minimizes the average squared distance between the original signal and its reduced linear representation.

3.5.2 Properties of the decomposition

In the following, properties of the proper orthogonal decomposition are listed. The proves of these statements are now classical and can be found in the literature.

— For a given a bounded domain, Hilbert-Schmidt theory applies and states that the eigenproblem has a denumerable set of solutions satisfying

$$\sum_{j=1}^{d} \int_{\Omega} R_{ij}(\boldsymbol{x}, \boldsymbol{x}')\, \phi_j^{(n)}(\boldsymbol{x}')\, \mathrm{d}\boldsymbol{x}' = \lambda^{(n)} \phi_i^{(n)}(\boldsymbol{x}) \tag{3.29}$$

where $\lambda^{(n)}$ and $\boldsymbol{\phi}_i^{(n)}$ represent respectively the eigenvalue and eigenfunction of order $n \geq 1$. Each eigenfunction is solution of the maximization problem (3.24) with the additional constraint of being orthogonal to all previous eigenfunctions.

— $\mathcal{R}$ can be shown to be self-adjoint and non negative, so that all eigenvalues are positive, real and converging, that is

$$\lambda^{(1)} \geq \lambda^{(2)} \geq \lambda^{(3)} \geq \ldots \geq 0, \quad \text{with} \quad \sum_{n=1}^{\infty} \lambda^{(n)} < +\infty. \tag{3.30}$$

— The set of eigenfunction $\boldsymbol{\phi}^{(n)}$ form a complete orthogonal set, meaning that almost every member of the set $\{\boldsymbol{d}(\boldsymbol{x}, t)\,,\ \boldsymbol{x} \in \Omega, t \in \mathbb{R}^+\}$ can be reconstructed as

$$\boldsymbol{d}(\boldsymbol{x}, t) = \sum_{n=1}^{\infty} \alpha^{(n)}(t)\, \boldsymbol{\phi}^{(n)}(\boldsymbol{x}) \tag{3.31}$$

where $\alpha^{(n)}$, projections of $\boldsymbol{d}$ onto $\boldsymbol{\phi}$, can be computed through the orthogonality

of the eigenfunctions $\boldsymbol{\phi}$ as

$$\alpha^{(n)}(t) = \langle \boldsymbol{d}, \boldsymbol{\phi} \rangle_{\mathcal{H}} = \sum_{i=1}^{d} \int_{\Omega} u_i(\boldsymbol{x}, t)\, \phi_i^{\dagger(n)}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}. \tag{3.32}$$

— Mercer's theorem: the two points correlation tensor R_{ij} can be written as a uniformly convergent series

$$R_{ij(\boldsymbol{x},\boldsymbol{x}')} = \sum_{n=1}^{\infty} \lambda^{(n)} \phi_i^{(n)}(\boldsymbol{x})\, \phi_j^{\dagger(n)}(\boldsymbol{x}'). \tag{3.33}$$

— Stemming from the diagonal representation of R_{ij}, the decomposition of $\boldsymbol{d}$ on the eigenfunctions $\boldsymbol{\phi}$ and their orthogonality, one has that

$$\overline{\alpha^{(n)}\alpha^{\dagger(m)}}^{X} = \delta_{nm}\lambda^{(n)}, \tag{3.34}$$

that means that the coefficients $\alpha^{(n)}$ are mutually uncorrelated and their mean square value are the eignevalues themselves.

— From Mercer's theorem and orthonormality of $\boldsymbol{\phi}^{(n)}$ one can write

$$\sum_{i=1}^{d} \int_{\Omega} R_{ij}(\boldsymbol{x}, \boldsymbol{x}')\, \mathrm{d}\boldsymbol{x} = \sum_{n=1}^{\infty} \lambda^{(n)} = E, \tag{3.35}$$

where E represents in the case of fluids with velocity field $\boldsymbol{d}$, the Turbulent Kinetic Energy (TKE) integrated over the domain Ω.

3.5.3 Algorithmic approach

Starting from the the formalism introduced in section 3.4, the experimental data is organized into a matrix $\boldsymbol{D}_{i,k}$ as explained in equation (3.13). The temporal correlation matrix $\boldsymbol{K} = \boldsymbol{D}^{\dagger}\boldsymbol{D} \in \mathbb{R}^{n_t \times n_t}$ is then computed as

$$\boldsymbol{K}_{ij} = \int_{\Omega} \boldsymbol{d}^{\dagger}(\boldsymbol{x}, t_i)\, \boldsymbol{d}(\boldsymbol{x}, t_j)\, \mathrm{d}\boldsymbol{x}. \tag{3.36}$$

The eigen-problem

$$\boldsymbol{K}\boldsymbol{\Psi} = \boldsymbol{\Psi}\boldsymbol{\Sigma} \tag{3.37}$$

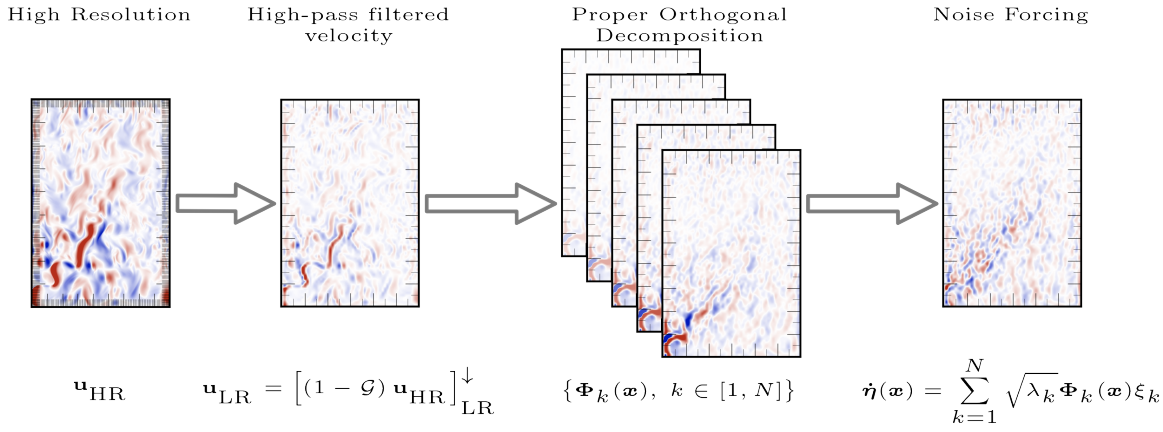


Figure 3.1 – Outline of the POD noise generation procedure

is then solved in order to define the temporal structures $\boldsymbol{\psi}(t)$. The corresponding spatial structures are then computed by projection of the data on these temporal structures, that means

$$\boldsymbol{\phi}_i(\boldsymbol{x}) = \frac{1}{T}\int_T \boldsymbol{d}(\boldsymbol{x}, t)\,\boldsymbol{\psi}_i(t)\,\mathrm{d}t \tag{3.38}$$

3.5.4 Noise ansatz

Employing POD on a set of velocity fluctuations of type (3.11) or (3.12) produces a set of velocity modes $\{\boldsymbol{\phi}_j(\boldsymbol{x}), \lambda_j, j = 1, \ldots, N\}$ that can be used to define the noise ansatz as

$$\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\mathbf{B}_t = \sqrt{\tau}\sum_{k=1}^{N}\lambda_k^{1/2}\boldsymbol{\phi}_k(\boldsymbol{x})\,\mathrm{d}\beta_t^k \tag{3.39}$$

with associated variance tensor computed as

$$\boldsymbol{a}(\boldsymbol{x}) = \tau\sum_{k=1}^{N}\lambda_k\boldsymbol{\phi}_k(\boldsymbol{x})\,\boldsymbol{\phi}_k^{\mathrm{T}}(\boldsymbol{x}). \tag{3.40}$$

If a non centred noise of type (2.11) is considered favourable, the time average $\overline{\boldsymbol{u}}^t$ that was removed from the initial data can be re-inserted in the simulation through the Girsanov correction, defining thus the noise as

$$\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\mathbf{B}_t = -\overline{\boldsymbol{u}}^t(\boldsymbol{x})\,\mathrm{d}t + \sqrt{\tau}\sum_{k=0}^{N}\lambda_k^{1/2}\boldsymbol{\phi}_k(\boldsymbol{x})\,\mathrm{d}\beta_t^k, \tag{3.41}$$

where $\boldsymbol{\sigma}_t \mathbf{Y}_t$ in (2.11) is defined as $\overline{\boldsymbol{u}}^t$.

3.6 Dynamical mode decomposition (DMD)

Dynamical Mode Decomposition is a technique that aims at providing informations about the dynamics of a flow. Assuming that a non-linear dynamical system of the kind

$$\frac{\mathrm{d}\boldsymbol{x}\left(\boldsymbol{x}_0, t\right)}{\mathrm{d}t} = f\left(\boldsymbol{x}\left(\boldsymbol{x}_0, t\right)\right), \tag{3.42}$$

this methodology, introduced by Schmid et al., 2008, constructs a proxy linear dynamical system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \mathcal{A}\boldsymbol{x} \tag{3.43}$$

with initial condition $\boldsymbol{x}\left(0\right)$ to describe the unknown non-linear dynamics. This linear system has a known analytical solution

$$\boldsymbol{x}\left(t\right) = \sum_{k=1}^{N} \boldsymbol{\phi}_k \exp\left(\mu_k t\right) b_k = \boldsymbol{\Phi} \exp\left(\boldsymbol{M} t\right) \boldsymbol{b}, \tag{3.44}$$

with $\boldsymbol{\phi}_k \in \mathbb{C}^d$ and $\mu_k \in \mathbb{C}$ the eigenvectors and eigenvalues (with corresponding matrices $\boldsymbol{\Phi}$ and $\boldsymbol{M}$) of the matrix $\mathcal{A}$, and $b_k \in \mathbb{C}$ the projection of the initial state $\boldsymbol{x}_0$ on the eigenvector basis (Sacco et al., 2020; Wu et al., 2021). The method identifies spatio-temporal coherent structures in high dimensional data, and it combines the favourable features of proper orthogonal decomposition in space and Fourier transform in time. The underlying mathematical formulation of this mode extraction from the time-resolved snapshots is closely related to the Arnoldi's algorithm (C. W. Rowley et al., 2009), with the advantage of being an equation-free procedure. In practice, DMD is based on a set of snapshots or measurements of a system in time, and it is completely data driven, with no additional assumptions on the underlying dynamics. Moreover, the DMD algorithm is tightly related to the spectral analysis of the Koopman operator, that is defined for every linear system (C. W. Rowley et al., 2009). Given Equation (3.44) it is possible to describe an analogous discrete-time system sampled at every Δt such as

$$\boldsymbol{x}_{k+1} = \mathbf{A}\boldsymbol{x}_k \tag{3.45}$$

with $\mathbf{A} = \exp(\mathcal{A}\Delta t)$, discrete map analogous to the continuous flow $\mathcal{A}$. The solution of this system may be expressed as

$$\boldsymbol{x}_k = \sum_{j=1}^{r} \boldsymbol{\phi}_j \lambda_j^k b_j = \boldsymbol{\Phi}\boldsymbol{\Lambda}^k \mathbf{b}, \tag{3.46}$$

where b_k is again the projection of the initial state on the eigenvector basis, $\boldsymbol{x}_1 = \boldsymbol{\Phi}\mathbf{b}$, with $\lambda_j = \exp(\mu_j t)$. The DMD involves approximating the eigendecomposition of the best-fit linear operator A that relates a state x_j at time t_j to the state x_{j+1} at the next timestep: The DMD algorithm is designed to produce a low-rank eigenvalue decomposition (3.46) of the matrix $\mathbf{A}$ that optimally fits the measured trajectory $\boldsymbol{x}_k$ for every $k = 1, 2, \ldots, m$ in a least square sense, so that

$$\|\boldsymbol{x}_{k+1} - \mathbf{A}\boldsymbol{x}_k\|_2 \tag{3.47}$$

is minimized across all points for $k = 1, 2, \ldots, m-1$.

3.6.1 Algorithmic approach

Algorithmically speaking, two matrices are defined as

$$\mathbf{D} = \begin{bmatrix} | & | & & | \\ \boldsymbol{d}_1 & \boldsymbol{d}_2 & \ldots & \boldsymbol{d}_{m-1} \\ | & | & & | \end{bmatrix}, \quad \mathbf{D}' = \begin{bmatrix} | & | & & | \\ \boldsymbol{d}'_1 & \boldsymbol{d}'_2 & \ldots & \boldsymbol{d}'_{m-1} \\ | & | & & | \end{bmatrix}, \tag{3.48}$$

where $\boldsymbol{d}'_k = \mathbf{F}(\boldsymbol{d}_k)$. Thanks to the local linear approximation one has that $\mathbf{D}' \approx \mathbf{A}\mathbf{D}$. The best linear fit to $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{D}'\mathbf{D}^{\ddagger} \tag{3.49}$$

where $\mathbf{D}^{\ddagger}$ is the Moore-Penrose pseudo-inverse. This solution minimizes the error defined as

$$\|\mathbf{D}' - \mathbf{A}\mathbf{D}\|_F, \tag{3.50}$$

$\|\cdot\|_F$ being the Frobenius norm. DMD circumvents the eigendecomposition of $\mathbf{A}$ by considering a rank-reduced representation in terms of a POD-projected matrix $\tilde{\mathbf{A}}$.

— Take the singular value decomposition (SVD) of $\mathbf{D}$

$$\mathbf{D} \approx \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\dagger}, \tag{3.51}$$

where † denotes the complex conjugate transpose, $\mathbf{U} \in \mathbb{C}^{n\times r}$, $\mathbf{\Sigma} \in \mathbb{C}^{r\times r}$, $\mathbf{V} \in \mathbb{C}^{m\times r}$ and r is the rank of the reduced SVD approximation to $\mathbf{D}$. The SVD reduction is exploited at this stage to perform a low-rank truncation of the data, meaning that if a low-dimensional structure is present in the data the singular values $\mathbf{\Sigma}$ will decrease sharply showing a limited number of dominant modes.

— The matrix $\mathbf{A}$ may be obtained by using the pseudoinverse of $\mathbf{D}$ obtained via SVD as

$$\mathbf{A} = \mathbf{D}'\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^{\dagger}, \tag{3.52}$$

but it is preferred to project the matrix $\mathbf{A}$ onto POD modes as

$$\tilde{\mathbf{A}} = \mathbf{U}^{\dagger}\mathbf{A}\mathbf{U} = \mathbf{U}^{\dagger}\mathbf{D}'\mathbf{V}\mathbf{\Sigma}^{-1}. \tag{3.53}$$

— Compute the eigendecomposition of $\tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\mathbf{\Lambda}, \tag{3.54}$$

where columns of $\mathbf{W}$ are eigenvectors and $\mathbf{\Lambda}$ is the diagonal matrix containing the eigenvalues λ_k.

— The eigendecomposition of $\mathbf{A}$ is reconstructed from $\mathbf{W}$ and $\mathbf{\Lambda}$. In particular, the eigenvalues of $\mathbf{A}$ are given by $\mathbf{\Lambda}$ and the eigenvectors of $\mathbf{A}$ are given by the columns of $\mathbf{\Phi}$:

$$\mathbf{\Phi} = \mathbf{D}'\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{W}. \tag{3.55}$$

— The values b_k can be computed as

$$\mathbf{b} = \mathbf{\Phi}^{\ddagger}\boldsymbol{x}_1. \tag{3.56}$$

The modes defined in (3.55) are called *exact DMD modes*, as it was proven in Tu et al. (2014) that these are the exact eigenvectors of the matrix $\mathbf{A}$.

3.6.2 Modes splitting (correlated vs. uncorrelated)

When the initial data are real valued fields, the eigenvectors, eigenvalues and amplitudes will be real or two-by-two complex conjugate, that is $\boldsymbol{\phi}_{2p} = \overline{\boldsymbol{\phi}_{2p+1}}$. Following the successful proposition of Li, Mémin, et al. (2022), we split the DMD modes into correlated and uncorrelated modes in order to define the Girsanov drift through the slow component

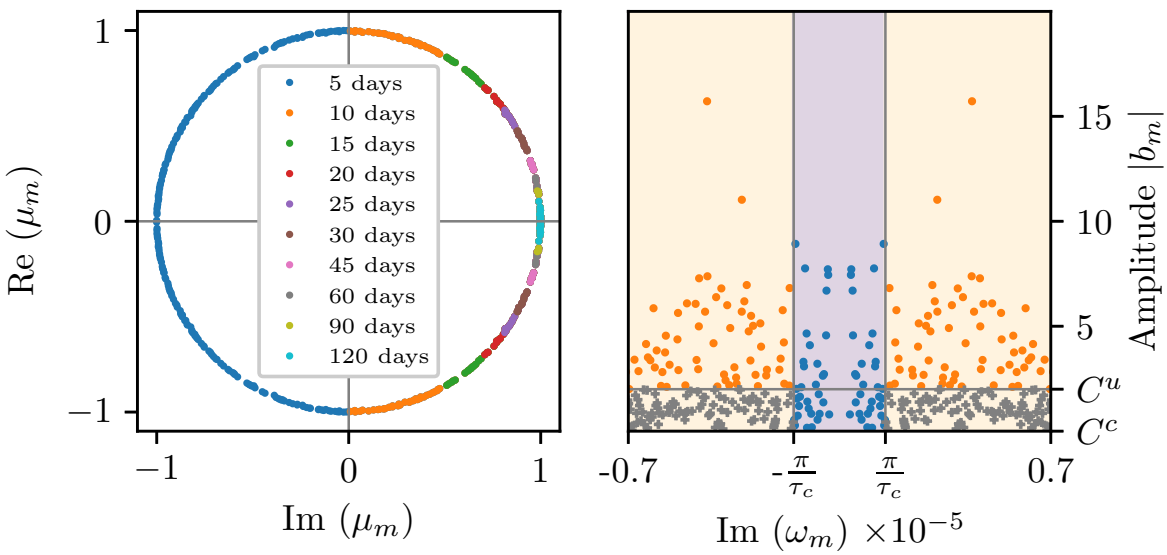


Figure 3.2 – Illustration of the selection of the DMD modes. On the left, frequencies of the modes are plotted on the unitary circle; they are coloured differently to represent their characteristic physical time scale. At this point, a threshold $\tau_c = 25$d is chosen to differentiate the correlated from the uncorrelated modes. On the right, over violet background are plotted the correlated modes, over orange background the uncorrelated modes. The amplitude threshold for the correlated mode C^c is set to zero, while for uncorrelated modes C^u is set to 2. The grey dots represent the set of uncorrelated modes below this threshold, that are thus discarded.

of the dynamics and the random noise through the fast component. The eigenvalue μ_m of Equation (3.44) can be split as $\mu_m = \sigma_m + \mathrm{i}\omega_m$, where the real part σ_m is the growth rate of the mode and ω_m is the periodic frequency of the mode m. The two sets of modes, $\mathcal{M}^u$ for the uncorrelated noise and $\mathcal{M}^c$ for the correlated part are defined as

$$\mathcal{M}^u = \left\{ m \in [1, N] \ : \ |\mu_m| \sim 1, \ |\omega_m| > \frac{\pi}{\tau_c}, \ |b_m| \geq C^u \right\}, \tag{3.57}$$

$$\mathcal{M}^c = \left\{ m \in [1, N] \ : \ |\mu_m| \sim 1, \ |\omega_m| \leq \frac{\pi}{\tau_c}, \ |b_m| \geq C^c \right\}, \tag{3.58}$$

where τ_c is a temporal separation scale between correlated and uncorrelated (usually set to a value for which a spectral gap is observed and fixed here to two months) and C^u, C^c are empirical cut-off of amplitudes. A visual representation of the aforementioned procedure is given in Figure 3.2.

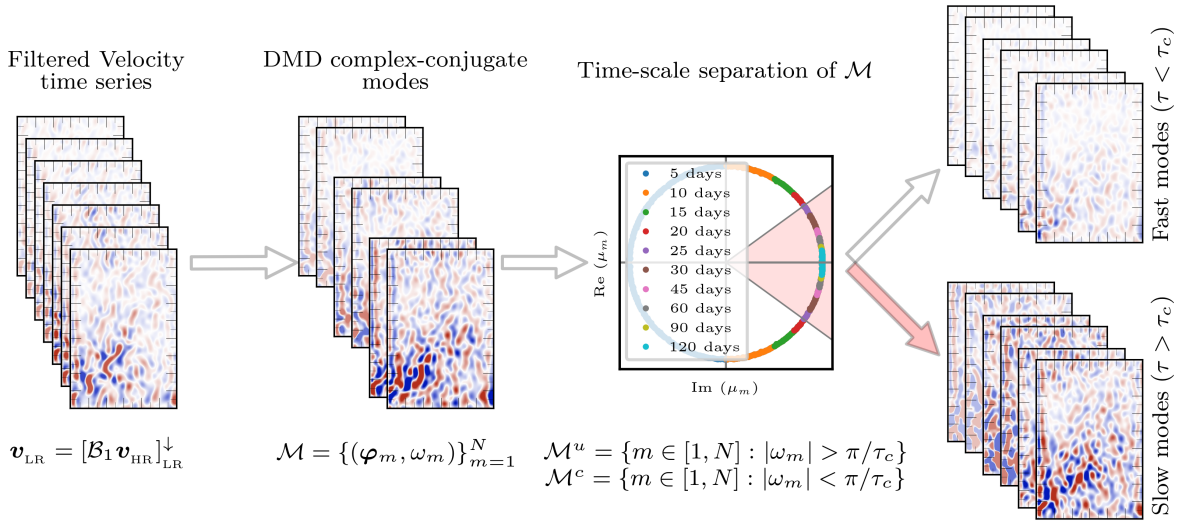


Figure 3.3 – Outline of the DMD noise generation procedure

3.6.3 Oblique projection onto of the modes

As the DMD modes are not orthogonal, a scaling is applied to avoid spurious effects and to make sure that the reconstructed data corresponds to an orthogonal projection onto the subspaces spanned by the set of modes contained in $\mathcal{M}^u$ and $\mathcal{M}^c$. The procedure reads as follow:

— Construct the Gramian matrix G of components $g_{m,n} = \langle \boldsymbol{\phi}_m, \boldsymbol{\phi}_n \rangle_{\mathcal{H}}$, with $m, n \in \mathcal{M}$;

— Invert the Gramian matrix and compute the dual set of modes $\boldsymbol{\Phi}^* = G^{-1}\boldsymbol{\Phi}$;

— Define the amplitudes as the initial state data on the dual set of modes: $\boldsymbol{\varphi}_m = \langle \boldsymbol{d}(\boldsymbol{x}, t_0), \boldsymbol{\phi}_m^* \rangle_{\mathcal{H}} \boldsymbol{\phi}_m$.

Such procedure is applied for $\mathcal{M} = \mathcal{M}^u$ and $\mathcal{M} = \mathcal{M}^c$ separately.

3.6.4 Noise ansatz

The dynamical mode decomposition is applied on a set of velocity fluctuations of type (3.11) or (3.12), so that a discrete-time proxy linear system is defined as

$$\boldsymbol{u}'(\boldsymbol{x}, t_{i+1}) \approx A\boldsymbol{u}'(\boldsymbol{x}, t_i). \tag{3.59}$$

Such (finite dimensional) linear dynamical system is known to have a general solution:

$$\boldsymbol{u}'(\boldsymbol{x},t) = \sum_{m=1}^{N} b_m \exp(\mu_m t)\, \boldsymbol{\phi}_m(\boldsymbol{x}), \tag{3.60}$$

where $\boldsymbol{\phi}_m(\boldsymbol{x}) \in \mathbb{C}^d$ are the eigenvectors of A associated to the eigenvalues $\mu_m \in \mathbb{C}$ and $b_m \in \mathbb{C}$ are amplitudes. In particular $\mu_m = \sigma_m + \mathrm{i}\omega_m$, the real part σ_m is the growth rate of the mode and ω_m is the periodic frequency of the mode m. The noise ansatz is thus

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t(\boldsymbol{x}) = \sqrt{\tau} \sum_{m \in \mathcal{M}^u} \exp(\mathrm{i}\omega_m t)\, \boldsymbol{\varphi}_m(\boldsymbol{x})\, \mathrm{d}\beta_m, \tag{3.61}$$

with associated stationary variance tensor

$$\boldsymbol{a}(\boldsymbol{x}) = \tau \sum_{m \in \mathcal{M}^u} \boldsymbol{\varphi}_m(\boldsymbol{x})\, \boldsymbol{\varphi}_m^{\dagger}(\boldsymbol{x}). \tag{3.62}$$

In Equations (3.61) and (3.62) τ is the process decorrelation time. Moreover, through the mode splitting procedure exposed in Section 3.6.2 the Girsanov drift can be naturally introduced in terms of the correlated, slowly varying component of the velocity fluctuations and time average as

$$\boldsymbol{\sigma}_t \mathbf{Y}_t(\boldsymbol{x}) = \overline{\boldsymbol{u}}^t(\boldsymbol{x}) + \sum_{m \in \mathcal{M}^c} \exp(\mathrm{i}\omega_m t)\, \boldsymbol{\varphi}_m(\boldsymbol{x}), \tag{3.63}$$

so that the Girsanov noise is defined as

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t(\boldsymbol{x}) = -\left[\overline{\boldsymbol{u}}^t(\boldsymbol{x}) + \sum_{m \in \mathcal{M}^c} \exp(\mathrm{i}\omega_m t)\, \boldsymbol{\varphi}_m(\boldsymbol{x})\right] \mathrm{d}t + \sqrt{\tau} \sum_{m \in \mathcal{M}^u} \exp(\mathrm{i}\omega_m t)\, \boldsymbol{\varphi}_m(\boldsymbol{x})\, \mathrm{d}\beta_m. \tag{3.64}$$

Each (eigen) frequency ω_n comes in pairs and each pair of complex Brownian motion are conjugates. The real and imaginary parts of the Brownian motion are independent. As such, both the noise and Girsanov drift are real-valued fields.

3.7 Pseudo-observations POD noise (PSO)

When dealing with eddy-resolving resolution simulation, a data driven approach might increase dramatically the memory footprint of the method, even if relying on order reduction procedures. For this reason, a data-free approach is worth exploring. The proposed

methodology relies on the application of a proper orthogonal decomposition to a set of *pseudo-observations* constructed from the current time velocity. A pseudo-observation of a given field at the point $\boldsymbol{x}$ is defined as the observation of that field in a second point $\boldsymbol{y}$ lying in a neighbourhood of $\boldsymbol{x}$. More formally, given two points, $\boldsymbol{x}$ and $\boldsymbol{y}$ such that $|\boldsymbol{x}-\boldsymbol{y}|<\delta$, one can assume that the observation of the field $\boldsymbol{u}$ at the two different points, $\boldsymbol{u}(\boldsymbol{x},t)$ and $\boldsymbol{u}(\boldsymbol{y},t)$, will be close in their statistics. A set of observations over a set of points $\{\boldsymbol{y}_i\}_{i=1}^N$ can therefore be considered as an ensemble of possible observations carrying the statistics of the point $\boldsymbol{x}$. This assumption can be considered as a form of self-similarity in space. We define the neighbourhood $\boldsymbol{x}$ as the set of points such that $|\boldsymbol{x}-\boldsymbol{y}|<\delta$ and denote it with $\nu(\boldsymbol{x})$, in this way the pseudo-observation of the velocity fluctuations can be constructed as

$$\tilde{\boldsymbol{u}}'(\boldsymbol{x},z,t)=\boldsymbol{u}(\boldsymbol{y},z,t)-\frac{1}{|\nu(\boldsymbol{x})|}\int_{\nu(\boldsymbol{x})}\boldsymbol{u}(\boldsymbol{\xi}-\boldsymbol{y},z,t)\,\mathrm{d}\boldsymbol{\xi} \tag{3.65}$$

where $\boldsymbol{y}\in\nu(\boldsymbol{x})$ is chosen uniformly in the neighbourhood, i.e. $\boldsymbol{y}\sim\mathcal{U}[\nu(\boldsymbol{x})]$. A set of N pseudo-observations for the velocity fluctuation is thus defined as $\{\tilde{\boldsymbol{u}}_i'(\boldsymbol{x},z,t)\}_{i=1}^N$ over the set of points $\{\boldsymbol{y}_i\}_{i=1}^N$. This procedure can be applied to all points of the domain being wary of sampling different of $\boldsymbol{y}_i$ at every $\boldsymbol{x}$, meaning that given two points, $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$, $\boldsymbol{y}_i^{(1)}$ and $\boldsymbol{y}_i^{(2)}$ are independent random variables. In the discrete case, given a field $\mathbf{u}(\mathbf{x},t)$, one constructs a set of pseudo-observations of the fluctuations $\{\tilde{\boldsymbol{u}}_i'(\boldsymbol{x},t)\}_{i=1}^N$ at each point $\boldsymbol{x}$ as illustrated before, where $\nu(\boldsymbol{x})$ is now the set of neighbouring points of $\boldsymbol{x}$, the integral is replaced with summation and the measure of ν is the integer measure.

3.7.1 Noise ansatz

The application of a proper orthogonal decomposition to the set of pseudo-observations $\{\tilde{\boldsymbol{u}}_i'(\boldsymbol{x},t)\}_{i=1}^N$ provides a set of orthogonal modes $\{\boldsymbol{\phi}_i\}_{i=1}^N$ that can be used to construct the noise as

$$\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t(\boldsymbol{x})=\sqrt{\tau}\sum_{k=1}^{N}\lambda_k^{1/2}(t)\,\boldsymbol{\phi}_k(\boldsymbol{x},t)\,\mathrm{d}\beta_k \tag{3.66}$$

and variance tensor defined as

$$\boldsymbol{a}(\boldsymbol{x},t)=\tau\sum_{k=1}^{N}\lambda_k\boldsymbol{\phi}_k(\boldsymbol{x},t)\,\boldsymbol{\phi}_k^{\mathrm{T}}(\boldsymbol{x},t)\,. \tag{3.67}$$

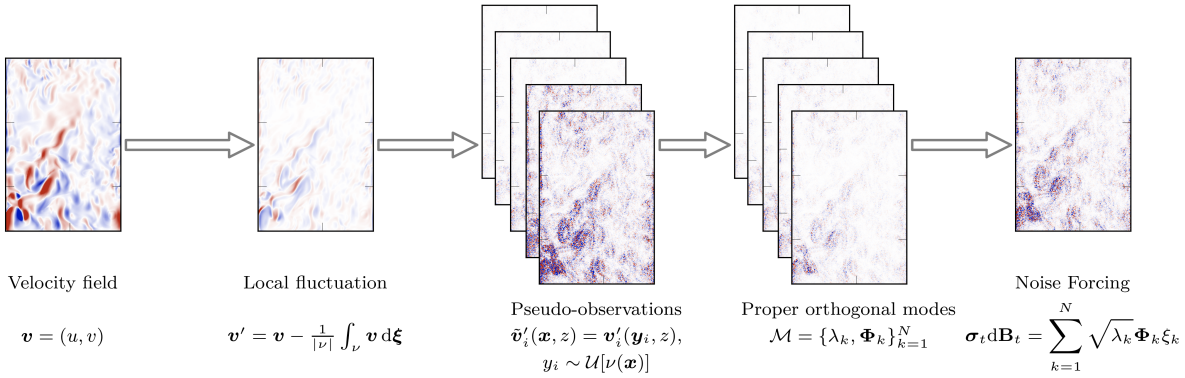


Figure 3.4 – Outline of the pseudo-observation noise generation procedure

The pseudo-observations generated with the aforementioned practice can be thought as corresponding to a virtual observation at a scale L, characteristic of the neighbourhood $\nu(\boldsymbol{x})$, and must be related to the simulation length scale ℓ. Harouna et al., 2017 proposes the scaling

$$\boldsymbol{a}_\ell = \left(\frac{\ell}{L}\right)^{2/3} \boldsymbol{a}_L, \tag{3.68}$$

based on the Kolmogorov-Richardson cascade assumption. A further manipulation leads to the definition of the scaling in terms of the integer patch size (*i.e.* the value N such that $L^2 = N\ell^2$) so to find

$$\boldsymbol{a}_\ell = \left(\sqrt{N}\right)^{2/3} \boldsymbol{a}_L. \tag{3.69}$$

Finally, the noise ansatz reads

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t(\boldsymbol{x}) = \sqrt{\tau}\left(\sqrt{N}\right)^{1/3} \sum_{k=1}^{N} \lambda_k^{1/2}(t)\,\boldsymbol{\phi}_k(\boldsymbol{x}, t)\,\mathrm{d}\beta_k \tag{3.70}$$

and variance tensor is defined as

$$\boldsymbol{a}(\boldsymbol{x}, t) = \tau\left(\sqrt{N}\right)^{2/3} \sum_{k=1}^{N} \lambda_k \boldsymbol{\phi}_k(\boldsymbol{x}, t)\,\boldsymbol{\phi}_k^{\mathrm{T}}(\boldsymbol{x}, t)\,. \tag{3.71}$$

3.8 Wavelet based noise (WLT)

A *wavelet* is a short-lived wave-like oscillation that is localized in time (or space). Wavelet processing has the characteristic of combining data processing in the time (or

space) domain and in the frequency domain, with a reasonable trade-off. The forward wavelet transform decompose the signal from the time (or space) domain to its representation in the wavelet basis, an oscillatory waveforms that reveal many signal properties and provide a sparse representation (Daubechies, 1992). Conversely, the inverse transform reconstruct the signal from its wavelet representation back to the time (spatial) domain. This operation is done through a function $\Phi(t)$, called *mother wavelet*, such that it is of zero mean (admissibility condition) and finite energy (localization condition), i.e.

$$\int_{\mathbb{R}} \Phi(t)\, \mathrm{d}t = 0 \quad \text{and} \quad \int_{\mathbb{R}} |\Phi(t)|^2\, \mathrm{d}t < +\infty. \tag{3.72}$$

This function is characterized further by two parameters, the shifting parameter s that translates the wavelet along the input signal and the scale parameter c that compresses or dilates the mother wavelet, so that the general form of Φ is:

$$\Phi_{s,c}(t) = \frac{1}{|s|^{1/2}} \Phi\left(\frac{t-\tau}{s}\right). \tag{3.73}$$

Chosen a scale s and a shifting τ, the value of the *wavelet coefficient* provides a quantification of the goodness of fit of the wavelet to the original signal, that is the scalar product between the original signal f and the scaled and translated wavelet

$$\langle f, \Phi_{s,c} \rangle (s,t) = \int_{\mathbb{R}} f(\tau) \frac{1}{|s|^{1/2}} \overline{\Phi}\left(\frac{t-\tau}{s}\right) \mathrm{d}\tau. \tag{3.74}$$

The set of wavelet coefficients is thus indexed by position in time (or space) and by the scale coefficient chosen. The inverse wavelet transform is defined as

$$f(t) = \int_{\mathbb{R}} \int_0^{\infty} \langle f, \Phi_{s,c} \rangle (s,\tau) \frac{1}{|s|^{1/2}} \Phi\left(\frac{t-\tau}{s}\right) \frac{\mathrm{d}s}{s^2}\, \mathrm{d}\tau, \tag{3.75}$$

that requires an integration on all the scales, expressing a zero information loss. Furthermore, in Fourier space the wavelet transform can be written as

$$\langle f, \Phi_{s,c} \rangle (s,t) = \sqrt{s} \int_{\mathbb{R}} \left[e^{2i\pi\omega t} \hat{\Phi}(s\omega)\, \hat{f}(\omega) \right] \mathrm{d}\omega \tag{3.76}$$

which characterizes the wavelet transform as a bandpass filter, with $\hat{\Phi}(s\omega)$ characterizing the shape of the filter. This formula also provides a mean of implementing the continuous wavelet transform.

3.8.1 Mathematical formulation of discrete wavelet transform (Multiresolution analysis)

Of all the possible implementations of the wavelet transform, the fast wavelet transform introduced by Mallat (1999) was chosen for its ease of implementation through filter banks and multiresolution analysis. The underlying idea is that of constructing approximations f_j of the function $f(t)$ at different scales j. The amount of information needed to go from a coarse approximation f_j to a finer approximation f_{j+1} is then described using orthogonal wavelets. Define V_j as the space of scale 2^j, meaning that it contains samples at a frequency 2^j. In other words, $V_j \subset L^2(\mathbb{R})$ is the set of functions f_j in $L^2(\mathbb{R})$ whose details are well represented at the scale 2^j. Assume then that there exists a function $\phi \in L^2(\mathbb{R})$ such that the family of scaled and shifted functions

$$\phi_{j,k}(t) = 2^{-j/2}\phi\left(2^{-j}t - k\right), \quad j,k \in \mathbb{Z} \tag{3.77}$$

is an orthonormal basis of V_j. This equation is the discretization of equation (3.73) with $s = 2^j$, for $j \in \mathbb{Z}$. The function ϕ is called the *scale function* of the multiresolution representation. A multiresolution representation in $L^2(\mathbb{R})$ is a sequence of closed subspaces V_j, $j \in \mathbb{Z}$, of $L^2(\mathbb{R})$, satisfying the following properties:

- $V_j \subset V_{j-1}$;
- $f \in V_j$ if and only if $f(2t) \in V_{j-1}$;
- $\bigcap_{j\in\mathbb{Z}} V_j = \emptyset$;
- $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$;
- there exists a function $\phi \in V_0$ such that the subspace V_0 is spanned by the orthonormal basis of the shifting $\{\phi(t-k)\,;\, k \in \mathbb{Z}\}$ of ϕ.

The first two conditions are of simple explanation, as increasing the j the sampling rate increases and the functions that were representable in the space scale V_{j-1} will also be representable in the scale space V_j, and this can be expressed also by the second condition, since scaling f of a factor of two will reduce the detail width of $1/2$, thus moving f to a finer subspace. This also means that all subspaces are scaled version of space V_0, of which ϕ and its shiftings are an orthonormal basis. Finally, the ambient set $L^2(\mathbb{R})$ contains all the possible scales, and the only function representable at all scales is a constant function, provided that this constant is zero due to the need of being square integrable. The representation operator for each scale space V_j is given by the orthogonal projection

over V_j, that is

$$\begin{aligned}\Pi_{V_j} &: L^2(\mathbb{R}) \to V_j \\ &f \to \Pi_{V_j}(f) = \sum_{n\in\mathbb{Z}} \langle f, \phi_{j,n}\rangle \phi_{j,n}\end{aligned}$$

that has also good convergence properties. From this definition it follows the characterization of the *detail space* W_j as the orthogonal complement of V_j, that is $V_{j-1} = V_j \oplus W_j$. Moreover, for every $j \neq k$ the two spaces W_j and W_k are orthogonal, so fixed an arbitrary index J_0 and for every $j < J_0$ one has

$$V_j = V_{J_0} \oplus \bigoplus_{k=0}^{J_0-j} W_{J_0-k} \tag{3.78}$$

and consequently $L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j$, i.e. the whole space can be spanned with the orthogonal basis of W_js. Here, we assume that a *wavelet* function ψ is contained in W_0 and that the set of shifted wavelets $\{\psi(t-k),\, k \in \mathbb{Z}\}$ spans W_0. It follows that the set of wavelets

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t - k\right), \quad j,k \in \mathbb{Z} \tag{3.79}$$

is an orthonormal basis of W_j, since every W_j inherits the scaling properties of V_j. The representation operator of a given function $f \in L^2(\mathbb{R})$ is given again by Galerkin projection

$$\begin{aligned}\Pi_{W_j} &: L^2(\mathbb{R}) \to W_j \\ &f \to \Pi_{W_j}(f) = \sum_{n\in\mathbb{Z}} \langle f, \psi_{j,n}\rangle \psi_{j,n}.\end{aligned}$$

It follows from (3.78) that the representation of a given function f on V_j can be written as

$$\Pi_{V_{j-J_0}}(f) = \Pi_{V_j}(f) + \Pi_{W_{j-1}}(f) + \ldots + \Pi_{W_{j-J_0}}(f) \tag{3.80}$$

that means that the projection over V_{j-J_0} is obtained summing the projections over the spaces W_{j-1} … W_{j-J_0} and the projection over V_j, that can be thought as a repeated projection over the spaces $V_{j-1}, \ldots, V_{j-J_0}$. The two functions ϕ and ψ can be related together through the so called *two-scale relation*. Considering ϕ the scale function of the multiresolution representation, one has that $\phi \in V_0 \subset V_{-1}$, so one can write ϕ as spanned

by the orthonormal basis ϕ_{-1} of V_{-1}, that is

$$\phi(t) = \sum_{k\in\mathbb{Z}} h_k \phi_{-1,k} = \sqrt{2} \sum_{k\in\mathbb{Z}} h_k \phi(2x-k) \tag{3.81}$$

with $h_k = \langle \phi, \phi_{-1,k} \rangle = \langle \phi(t), 2\phi(2t-k) \rangle$. Similarly, the wavelet ψ belongs to W_0 and, since $V_{-1} = V_1 \oplus W_0$, $\psi \in V_{-1}$, therefore

$$\psi(t) = \sum_{k\in\mathbb{Z}} g_k \phi_{-1,k} = \sqrt{2} \sum_{k\in\mathbb{Z}} g_k \phi(2x-k) \tag{3.82}$$

where $g_k = \langle \psi, \phi_{-1,k} \rangle = \langle \psi(t), 2\phi(2t-k) \rangle$. Functions ϕ and ψ verify the following relation where sequences $h_k = \langle \phi(t), 2\phi(2t-k) \rangle$ and $g_k = \langle \psi(t), 2\phi(2t-k) \rangle$ are called *conjugate mirror filters*. Consider a subspace V_0 and the ladder of nested subspaces $\ldots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset \ldots$, with all V_j scaled version of the V_0. One can define the difference subspace W_j as the orthogonal complement of V_j and V_{j-1}, that is $V_j = V_{j+1} \oplus W_{j+1}$. With this one can also represent the space of square integrable functions over the real line as

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j \tag{3.83}$$

and thus represent any function $f \in L^2(\mathbb{R})$ as the sum of its projections over each W_j, that is

$$f = \sum_{j\in\mathbb{Z}} \Pi_{W_j}(f) \tag{3.84}$$

where $\Pi_{W_j}(\cdot)$ is the orthogonal projection over the space W_j. From the definition of the space W_j one can however express every function $f_j \in V_j$ as

$$f_j = \Pi_{V_{j+1}}(f) + \Pi_{W_{j+1}}(f). \tag{3.85}$$

Define $\{\phi_{j,n};\, n \in \mathbb{Z}\}$ as the set of basis functions for V_j and $\{\psi_{j,n};\, n \in \mathbb{Z}\}$ as the set of basis functions for W_j, then the projection operators can be represented as

$$\Pi_{V_j}(f) = \sum_{n\in\mathbb{Z}} \langle f, \phi_{j,n} \rangle \phi_{j,n} = \sum_n \left(\int_{\mathbb{R}} f(t)\, \overline{\phi_{j,n}(t)}\, \mathrm{d}t \right) \phi_{j,n} \tag{3.86}$$

$$\Pi_{W_j}(f) = \sum_{n\in\mathbb{Z}} \langle f, \psi_{j,n} \rangle \psi_{j,n} = \sum_n \left(\int_{\mathbb{R}} f(t)\, \overline{\psi_{j,n}(t)}\, \mathrm{d}t \right) \psi_{j,n} \tag{3.87}$$

Wavelet *analysis* is now provided by the recursive application of this multiresolution concept as a function f_j is represented as the sum of the projections over the spaces $W_{j+1}, \ldots, W_{j+N}$ and a residual given by V_{j+N} as

$$f_j = \Pi_{V_{j+N}}(f) + \Pi_{W_{j+N}}(f) + \ldots + \Pi_{W_{j+2}}(f) + \Pi_{W_{j+1}}(f). \tag{3.88}$$

Wavelet *synthesis* can be also deduced from this formula, as $W_j \subset V_{j-1}$ and $V_j \subset V_{j-1}$ and thus one can reconstruct the function by proceeding backwards with no loss of information.

$$w(t) = \sum_{k=0}^{2^C-1} \langle w, \phi_{C,k} \rangle_{L^2} \phi_{C,k}(t) + \sum_{j=C}^{F} \sum_{k=0}^{2^j-1} \langle w, \psi_{j,k} \rangle_{L^2} \phi_{j,k}(t) \tag{3.89}$$

In the previous equation $\{\phi_{C,k}\}_k$ and $\{\psi_{j,k}\}_k$ are orthonormal bases of V_C and W_j respectively and they are defined by dilatations and translations of the so called *scale function* ϕ and corresponding *wavelet function* ψ.

3.8.2 Algorithmic approach

Algorithmically speaking, the application of the direct wavelet transform boils down to a matrix vector product with the matrix $\mathbf{W}$ defined as

$$\mathbf{W} = \begin{bmatrix} h_d & h_{d-1} & \ldots & h_2 & h_1 & 0 & & \ldots & & 0 & 0 \\ g_d & g_{d-1} & \ldots & g_2 & g_1 & 0 & & \ldots & & 0 & 0 \\ 0 & 0 & h_d & h_{d-1} & \ldots & h_2 & h_1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & g_d & g_{d-1} & \ldots & g_2 & g_1 & 0 & \ldots & 0 & 0 \\ \vdots & & & & & & & & & \vdots & \vdots \\ h_{d-2} & \ldots & h_1 & 0 & & & \ldots & & & h_d & h_{d-1} \\ g_{d-2} & \ldots & g_1 & 0 & & & \ldots & & & g_d & g_{d-1} \end{bmatrix}, \tag{3.90}$$

and the input vector $\mathbf{x}$. The application of the inverse wavelet transform is the matrix-vector product of the transformed vector $\mathbf{y}$ and the inverse of $\mathbf{W}$, that is $\mathbf{W}^{-1} = \mathbf{W}^{\mathrm{T}}$, being the wavelet family orthogonal and so the matrix $\mathbf{W}$. Two aspects of the construction of the matrix $\mathbf{W}$ must be looked up closely: the row-wise alternation between high-pass and low pass filtering and the circulant nature of $\mathbf{W}$ in the last two rows. This latter aspect can be avoided by extending the matrix instead of wrapping it. Regarding the former aspect, care is usually put to separate the result of the two filters in two different parts

of the output vector, introducing a complexity in the algorithm given by the necessity of storing the data in the correct place and in the selection of the correct data in the reconstruction. These two aspects are crucial when designing a wavelet filter operating on spatial data that cannot be accessed completely but are split into sub-domains in a domain decomposition approach. In particular, if physically meaningful, wrapping of data is allowed only on global boundaries, meaning that interior subdomains should access their neighbouring domains data rather than wrapping the data that they can locally access. Furthermore the filtering and decimation strategy of the multi-resolution analysis is difficult to port to data that do not come in power-of-2 length. These aspects where closely studied in order to implement correctly the wavelet filter bank.

Algorithm 1 Multilevel wavelet algorithm

```
procedure WAVELET(x,y,n,l)
    s = 0
    if direct transform then                          ▷ Direct wavelet transform
        m = n
        while s < l do
            y(1:m) = SINGLE_LEVEL_WAVELET(x, W, "dir")
            m = m/2
            s = s + 1
        end while
    else if Inverse transform then                    ▷ Inverse wavelet transform
        m = n/2^l
        while s < l do
            x(1:m) = SINGLE_LEVEL_WAVELET(y, W, "inv")
            m = m * 2
            s = s + 1
        end while
    end if
end procedure
```

3.8.3 Noise ansatz

The wavelet analysis is applied to a snapshot of the velocity fluctuations, $\boldsymbol{u}'$, computed with respect to a local average of type (3.7). The result of this operation is a set of *details* $\langle \boldsymbol{u}', \boldsymbol{\psi}_{C,k} \rangle_{L^2}$ and a large scale component $\langle \boldsymbol{u}', \boldsymbol{\phi}_{C,k} \rangle_{L^2}$. These fields are then randomised with a set of Brownian motions β_t defined on the wavelet coefficients grid, so that the

Algorithm 2 Single level wavelet algorithm

```
function SINGLE_LEVEL_WAVELET(sig,W,direction)
    m = len(sig)
    p = len(W) - 1
    q = ( p - 1 ) / 2
    if direction = dir then                                  ▷ Direct wavelet transform
        i = 1                                                ▷ Position in the output vector
        initialize working array z up to m
        for j = 1, m - 1, 2 do                               ▷ Loop on the input vector
            for k = 0, p - 1, 2 do                           ▷ Loop on convolution filter weights
                j0 = i4_wrap ( j + k, 1, m )
                j1 = i4_wrap ( j + k + 1, 1, m )
                z(i) = z(i) + W( k) * sig(j0) + W( k+1) * sig(j1)
                z(i+m/2) = z(i+m/2) + W(p-k) * sig(j0) - W(p-k-1) * sig(j1)
            end for
            i = i + 1                                        ▷ Advance in the output vector
        end for
     return y = z
    else if direction = inv then                             ▷ Inverse wavelet transform
        j = 1                                                ▷ Position in the output vector
        initialize working array z up to m
        for i = - q + 1, m / 2 - q do                        ▷ Loop on the input vector
            for k = 0, p - 1, 2 do                           ▷ Loop on convolution filter weights
                i0 = i4_wrap ( i + k / 2, 1, m / 2 )
                i1 = i4_wrap ( i + m / 2 + k / 2, m / 2 + 1, m )
                z(j) = z(j) + W(p-k-1) * sig(i0) + W(k+1) * sig(i1)
                z(j+1) = z(j+1) + W(p-k) * sig(i0) - W(k) * sig(i1)
            end for
            j = j + 1                                        ▷ Advance in the output vector
        end for
     return x = z
    end if
end function
```

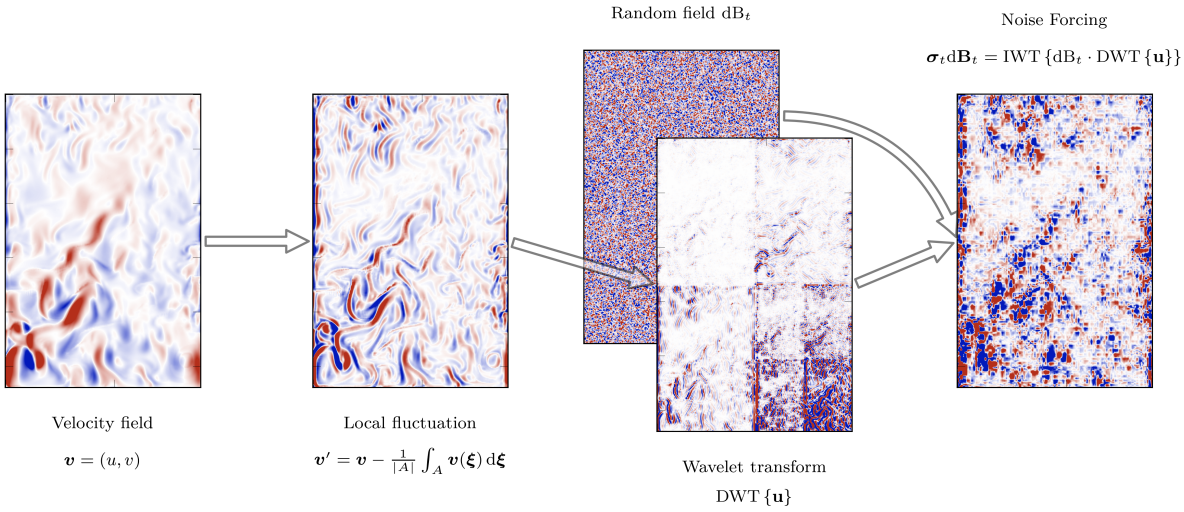


Figure 3.5 – Outline of the wavelet noise generation procedure

noise wavelet ansatz can be defined as

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t(\boldsymbol{x}) = \sum_{k=0}^{2^C-1} \langle \boldsymbol{u}'_{(t)}, \boldsymbol{\phi}_{C,k} \rangle_{L^2} \mathrm{d}\beta_t^{C,k} \boldsymbol{\phi}_{C,k}(\boldsymbol{x}) + \sum_{j=C}^{F} \sum_{k=0}^{2^j-1} \langle \boldsymbol{u}'_{(t)}, \boldsymbol{\psi}_{j,k} \rangle_{L^2} \mathrm{d}\beta_t^{j,k} \boldsymbol{\phi}_{j,k}(\boldsymbol{x}). \quad (3.91)$$

The subscript (t) emphasizes that the wavelet processing is applied to the current-state n of the simulation. The first component of the noise represents the randomised large scale dynamics, and is set to zero to represent the small scale features only and perform a spatial Reynolds-like decomposition. The definition of the variance tensor can then be based on the definition of the *details*. Such type of noise terms can easily be shown to be well defined. It is spatially regular and its regularity is given by the choice of the wavelet basis. The wavelet transform conveys a natural multi-resolution structure to the noise as well as a natural notion of spatial scale at each level of the multi-resolution hierarchy. Alternatively, the wavelet analysis is applied to a snapshot of the velocity $\boldsymbol{u}$ and then synthesis is applied only to the set of details, randomized with the set of Brownian motions β_t, that is

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t(\boldsymbol{x}) = \sum_{j=C}^{F} \sum_{k=0}^{2^j-1} \langle \boldsymbol{u}_{(t)}, \boldsymbol{\psi}_{j,k} \rangle_{L^2} \mathrm{d}\beta_t^{j,k} \boldsymbol{\phi}_{j,k}(\boldsymbol{x}). \quad (3.92)$$

as in this case the spatial filtering is done by removing the large scale component recognised by the wavelet analysis instead that with a filtering of type (3.7). The definition of the variance tensor in this case remains open to discussion.

3.9 Vertical profile prescription noise (VPP)

An efficient way to construct a noise that does not rely on data is that of applying directly the definition of the noise as given in Equation (3.1),

$$\sigma_t \mathrm{dB}_t^i(\boldsymbol{x}, z) = c(\boldsymbol{x}, z) \int_\Omega \breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{y}, t)\, \mathrm{dB}_t^k(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}, \quad i, k = 1, 2. \tag{3.93}$$

In the previous equation, it has to be remarked that $\boldsymbol{x}$ and $\boldsymbol{y}$ are only horizontal. The function $c(\boldsymbol{x}, z)$ is a vertical profile that is imposed on the noise to provide to it a connection with the current state of the flow. The integral kernel $\breve{\sigma}$ is chosen to have a simple Gaussian structure on the horizontal, such as

$$\breve{\sigma}_{ik}(\boldsymbol{x}, \boldsymbol{y}, t) = \frac{1}{\sqrt{2\pi} d} \exp\left[-\frac{|\boldsymbol{x} - \boldsymbol{y}|^2}{2d^2}\right] \coloneqq \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y}), \quad \forall i, k = 1, 2, \tag{3.94}$$

of a given width d. This filter is applied to a set of Brownian motions $\mathrm{d}\mathbf{B}_t$ defined on the computational grid. The associated variance tensor $\boldsymbol{a}$ is diagonal and defined as

$$\boldsymbol{a}(\boldsymbol{x}, z) = \begin{bmatrix} \alpha(\boldsymbol{x}, z) & 0 & 0 \\ 0 & \alpha(\boldsymbol{x}, z) & 0 \\ 0 & 0 & \beta(\boldsymbol{x}, z) \end{bmatrix}, \tag{3.95}$$

where

$$\alpha(\boldsymbol{x}, z) = c^2(\boldsymbol{x}, z)\, \delta_{ij} \mathrm{d}t \int_\Omega \breve{\sigma}_d^2(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{y},$$
$$\beta(\boldsymbol{x}, z) = \sum_{n=1}^{k} h_n c^2(\boldsymbol{x}, n)\, \mathrm{d}t \int_\Omega \left(\partial_i \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\right)^2 \mathrm{d}\boldsymbol{y}.$$

The vertical profile can follow any shape. In particular, the chosen approach is that of computing a normalised profile for the vertical kinetic energy as

$$c(\boldsymbol{x}, z) = \frac{1}{\|c(\boldsymbol{x}, z)\|} \frac{1}{\eta - \eta_b} \int_{\eta_b}^{z} \left[\frac{1}{|\nu(\boldsymbol{x})|} \int_{\nu(\boldsymbol{x})} w(\boldsymbol{\xi}, t)\, \mathrm{d}\boldsymbol{\xi}\right]^2 \mathrm{d}\zeta, \tag{3.96}$$

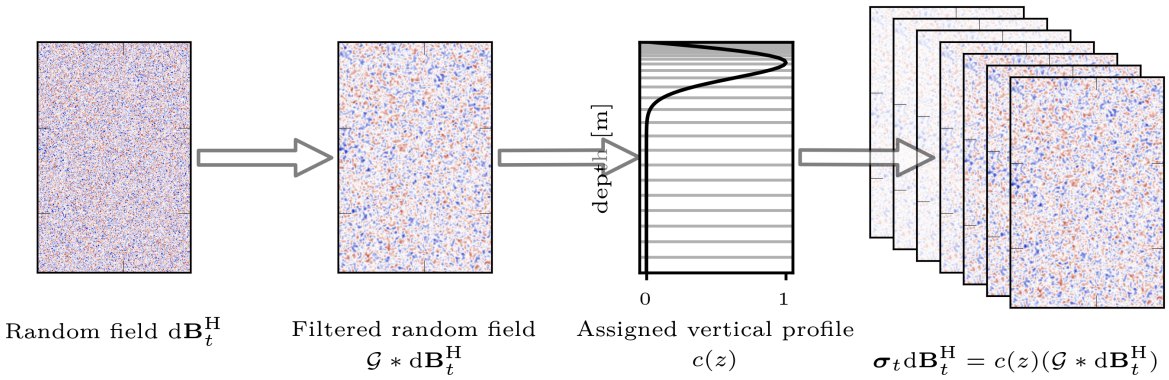


Figure 3.6 – Outline of the VPP noise generation procedure

where the spatial average increases the stability of the method by removing too small variations.

Explicit definition of the quadratic variation

In this section the explicit computation of the quadratic covariation terms is performed. The term $\boldsymbol{\sigma}\mathrm{d}\mathbf{B}_t^{i,k}$ will represent the $i-$th component of the noise at level k. In the following, the system has been discretized in the vertical for convenience, substituting the continuous formulation (3.8) with its discrete counterpart. The purely horizontal terms read:

$$\begin{aligned}
\langle \boldsymbol{\sigma}\mathrm{d}\mathbf{B}^{i,k}, \boldsymbol{\sigma}\mathrm{d}\mathbf{B}^{j,k}\rangle_t &= c^2(\boldsymbol{x}, z)\, \langle \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{dB}^{i,k}(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}, \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y}')\, \mathrm{dB}^{j,k}(\boldsymbol{y}')\, \mathrm{d}\boldsymbol{y}'\rangle_t \\
&= c^2(\boldsymbol{x}, z) \int_\Omega \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\, \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y}') \underbrace{\mathrm{d}\langle \mathrm{B}^{i,k}(\boldsymbol{y}), \mathrm{B}^{j,k}(\boldsymbol{y}')\rangle_t}_{\delta(\boldsymbol{y}-\boldsymbol{y}')\delta_{ij}\mathrm{d}t} \mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{y}' \\
&= c^2(\boldsymbol{x}, z)\, \delta_{ij}\mathrm{d}t \int_\Omega \breve{\sigma}_d^2(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{y} \coloneqq \alpha(\boldsymbol{x}, z)\, \mathrm{d}t.
\end{aligned}$$

Mixed horizontal-vertical terms read:

$$\begin{aligned}
\langle \boldsymbol{\sigma}\mathrm{d}\mathbf{B}^{i,k}, \boldsymbol{\sigma}\mathrm{d}\mathbf{B}^{z,k}\rangle_t &= \langle c(\boldsymbol{x}, k) \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{dB}^{i,k}(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}, -\sum_{\ell=1}^{k} h_\ell c(\boldsymbol{x}, \ell)\, \partial_j \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y}')\, \mathrm{dB}^{j,k}(\boldsymbol{y}')\, \mathrm{d}\boldsymbol{y}'\rangle_t \\
&= -\sum_{\ell=1}^{k} h_\ell c^2(\boldsymbol{x}, \ell) \int_\Omega \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\, \partial_j \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y}') \underbrace{\mathrm{d}\langle \mathrm{B}^{i,k}(\boldsymbol{y}), \mathrm{B}^{j,\ell}(\boldsymbol{y}')\rangle_t}_{\delta(\boldsymbol{y}-\boldsymbol{y}')\delta_{ij}\delta_{k\ell}\mathrm{d}t} \mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{y}' \\
&= -h_k c^2(\boldsymbol{x}, k)\, \mathrm{d}t \int_\Omega \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\, \partial_i \breve{\sigma}_d(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{y} = 0,
\end{aligned}$$

where the last result stems from the fact that integrating a symmetric function as $\breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y})$ against an antisymmetric function as $\partial_i\breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y})$ should give zero as a result. Finally, the last term reads

$$\begin{aligned}
\langle \boldsymbol{\sigma}\mathrm{d}\mathbf{B}^{z,k}, \boldsymbol{\sigma}\mathrm{d}\mathbf{B}^{z,k}\rangle_t &= \langle -\sum_{n=1}^{k} h_n c(\boldsymbol{x},n)\,\partial_i \int_\Omega \breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y})\,\mathrm{dB}^{i,n}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, -\sum_{m=1}^{k} h_m c(\boldsymbol{x},m)\,\partial_j \int_\Omega \breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y}')\,\mathrm{dB}^{j,m}(\boldsymbol{y}')\,\mathrm{d}\boldsymbol{y}'\rangle_t \\
&= -\sum_{n=1}^{k}\sum_{m=1}^{k} h_n h_m c(\boldsymbol{x},n)\,c(\boldsymbol{x},m) \int_\Omega\int_\Omega \partial_i\breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y})\,\partial_j\breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y}')\,\underbrace{\mathrm{d}\langle \mathrm{B}^{i,n}(\boldsymbol{y}),\mathrm{B}^{j,m}(\boldsymbol{y}')\rangle_t}_{\delta(\boldsymbol{y}-\boldsymbol{y}')\delta_{ij}\delta_{nm}\mathrm{d}t}\,\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{y}' \\
&= \sum_{n=1}^{k} h_n c^2(\boldsymbol{x},n)\,\mathrm{d}t \int_\Omega \left(\partial_i\breve{\sigma}_d(\boldsymbol{x},\boldsymbol{y})\right)^2 \mathrm{d}\boldsymbol{y} \coloneqq \beta(\boldsymbol{x},z)\,\mathrm{d}t.
\end{aligned}$$

3.10 Isopycnal projection

The noise models presented so far can be classified in two main classes: offline data models and online data models. The latter use available data from the current simulation to tie the noise to the solution itself. The former, offline data models, are suitable for introducing an external information in a simulation, but this information is not tied to the current state and thus it is poses a risk and a challenge. To provide a link between the external data and the current simulation the idea of constraining the noise to live onto isopycnal surfaces is borrowed from classical approaches (see the short discussion given in Section 1.4). After construction with the offline data through Equations (3.39) or (3.64), the noise $\boldsymbol{\sigma}_{t,\theta}\mathrm{d}\mathbf{B}_t$ is constrained to live on the tangent space of the isopycnal surfaces. This procedure is operationally implemented as the application of an isopycnal projection operator P^ρ

$$\mathrm{P}^\rho = I - \frac{\nabla\rho\,(\nabla\rho)^{\mathrm{T}}}{|\nabla\rho|^2} \tag{3.97}$$

to the noise. Being the density function of temperature and salinity, $\rho = \rho(T,S,z)$, the isopycnal projection operator carries information about the current state of the simulation. The projected noise $\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^\rho(\mathbf{x}) = \mathrm{P}^\rho\,\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t(\mathbf{x})$ is thus strongly tied to the evolution of the flow density. This technique has also the strong benefit of stabilising the noise, so that a higher noise level and thus a higher energy level can be used in the simulation.

Chapter 4

L'ocean de bon secours

Abstract

This chapter presents the results obtained from the implementation of the previously discussed methods in the ocean dynamics core NEMO. The name of this set of experiments has been jovially inspired by Saint-Malo's attraction, *la piscine de bon secours.*

The experiments are performed using the state-of-the-art modelling framework for oceanography and climatology, NEMO, that stands for "Nucleus for European Modelling of the Ocean" (Madec et al., 2019). The "blue ocean" engine of NEMO solves the primitive equations of ocean thermodynamics on a curvilinear orthogonal Arakawa C-grid with full or partial step s-coordinate, allowing the usage of z- or σ- coordinates. Prognostic variables are the horizontal velocity field, a linear or non-linear sea surface height, the conservative temperature and the absolute salinity. Non-linear sea surface height is discretized in time with a three-step Generalized Forward Backward algorithm based on third order Adams-Bashford and fourth order Adams-Moulton schemes, as proposed by Shchepetkin et al. (2009).

The domain configuration is a rectangular basin on the β-plane centred at $\sim 30°$N, rotated by 45° degrees, 3180 km long, 2120 km wide and 4 km deep, bounded by vertical walls and by a flat bottom. Seasonal winds (surface wind stress τ) and buoyancy changes (freshwater flux F, restoring towards an apparent air temperature $T^\star$ and penetrative heat Q) are imposed as external forcings to induce the creation of a strong jet that separates a cold sub-polar gyre from a warm sub-tropical gyre. This jet starts at the most eastward point of the domain and moves towards the interior of the domain. A recirculation gyre is as well visible in the southern corner. The complete details of this configuration are given in Lévy, Klein, et al. (2010) and Lévy, Resplandy, et al. (2012), while Figure 4.1 shows the geographical collocation of the basin, the applied forcings and provides a

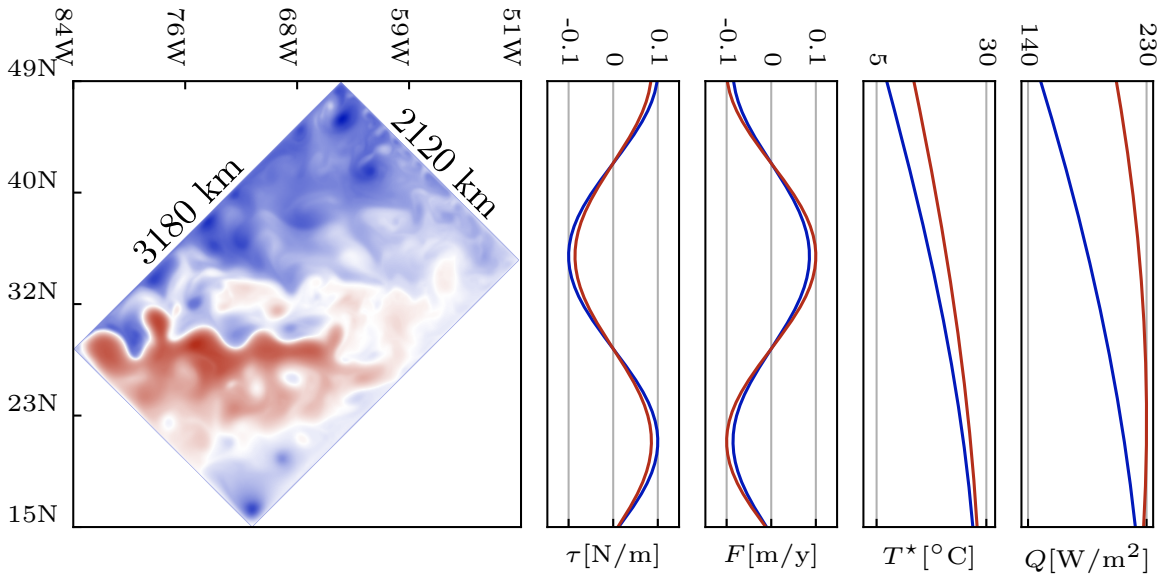


Figure 4.1 – Geographical collocation of the domain with its external forcings: surface wind stress τ imposed in the momentum equations, freshwater flux F imposed in the sea surface equation, restoring towards an apparent air temperature $T^\star$ and penetrative heat Q in the temperature equation. Inside the domain, sea surface height is plotted, showing the anomaly caused by the jet stream and thus providing its localization.

first glimpse of the jet stream in terms of sea surface height. The vertical domain, for all experiments, is discretised into 30 levels of increasing thickness with depth. Close to the ocean surface, the thickness is ∼10m and it increases to ∼500m at the ocean floor. The vertical discretization is kept constant as an increase was observed to have minimal effect by Lévy, Klein, et al. (2010). Experiments at varying horizontal resolutions (R27, R9, R3) are performed. The R3 is a coarse simulation with a horizontal resolution of 35.3 km (1/3°) which corresponds roughly to a central value of mesoscale eddy-permitting resolution. The finer resolution at R9 of 11.8 km (1/9°) is approaching realistic eddy-resolving ocean models (Sasaki et al., 2008; Maltrud et al., 2005). The finest resolution at R27 with 3.9 km (1/27°) is at a resolution where a significant portion of sub-mesoscale eddies would be resolved. The choice of resolutions correspond to configurations that are well documented in literature (Lévy, Klein, et al., 2010; Lévy, Resplandy, et al., 2012) and for which values of the dissipation coefficient have been established. Additionally, the choice of resolutions was dictated by major differences observed in the classical determin-

Table 4.1 – Model parameters for different horizontal resolutions

	R3	R9	R27
Horizontal resolution (in °)	1/3°	1/9°	1/27°
Horizontal resolution (in km)	35.3 km	11.8 km	3.9 km
Mesh grid (Nx × Ny)	90 × 60	270 × 180	820 × 540
Vertical levels (Nz)	30	30	30
Time step (in s)	3600	1200	300
Eddy Viscosity (κ_{M})	$-10^{12}\mathrm{m}^4 s^{-1}$	$-5 \times 10^{10}\mathrm{m}^4\mathrm{s}^{-1}$	$-5 \times 10^{9}\mathrm{m}^4\mathrm{s}^{-1}$
Eddy diffusivity (κ_{T})	$300\,\mathrm{m}^2\mathrm{s}^{-1}$	$-5 \times 10^{10}\mathrm{m}^4\mathrm{s}^{-1}$	$-10^{9}\mathrm{m}^4\mathrm{s}^{-1}$

istic simulation when the resolution is coarsened from R9 to R3 (Lévy, Resplandy, et al., 2012). A fourth high resolution experiment is performed, R54, with a resolution of 1.85 km (1/54°). This simulation has been set up following Lévy, Klein, et al. (2010) and Lévy, Resplandy, et al. (2012). It is only used in the next section for a qualitative description of the deterministic scenarios and not for direct comparison nor to provide synthetic data for data-driven noise models. For this reason no details about this simulation are provided in this work. The first run is initialized with a 1/3° resolution simulation, spun up for 1000 years. The resulting state is then interpolated on the 1/9° and 1/27° grids to initialize the runs. The R27d simulation has been spun-up for 100 years before collecting data for the LU framework. Similarly, an additional 1/9° deterministic simulation has been spun up for 100 years in similar conditions in order to construct an initial state for the deterministic and stochastic 1/3° simulations. Three different initial conditions are then obtained, corresponding to a 1100 years spin up, and used for the operational runs discussed in the following. In order to assess the benefits brought by this stochastic approach, each stochastic simulation is compared to its deterministic counterpart at the same resolution and at highest resolution (1/27°). The model parameters for the three resolutions are given in Table 4.1. A bi-harmonic horizontal momentum diffusion (characterised by κ_{M}) is implemented for all the resolutions. For the tracers, i.e. temperature and salinity, diffusion is implemented along isopycnal surfaces without horizontal background for R3 while for the finer resolution a bi-harmonic horizontal diffusion (characterised by κ_{T}) is implemented. The values for the diffusion/friction coefficients are obtained from Lévy, Klein, et al. (2010) and Lévy, Resplandy, et al. (2012). Vertical diffusion for all resolutions are implemented via the turbulent closure model developed by Bougeault et al. (1989). The model is implemented with a background value of $10^{-5}\mathrm{m}^2\mathrm{s}^{-1}$. Partial slip boundary conditions are implemented in all the experiments.

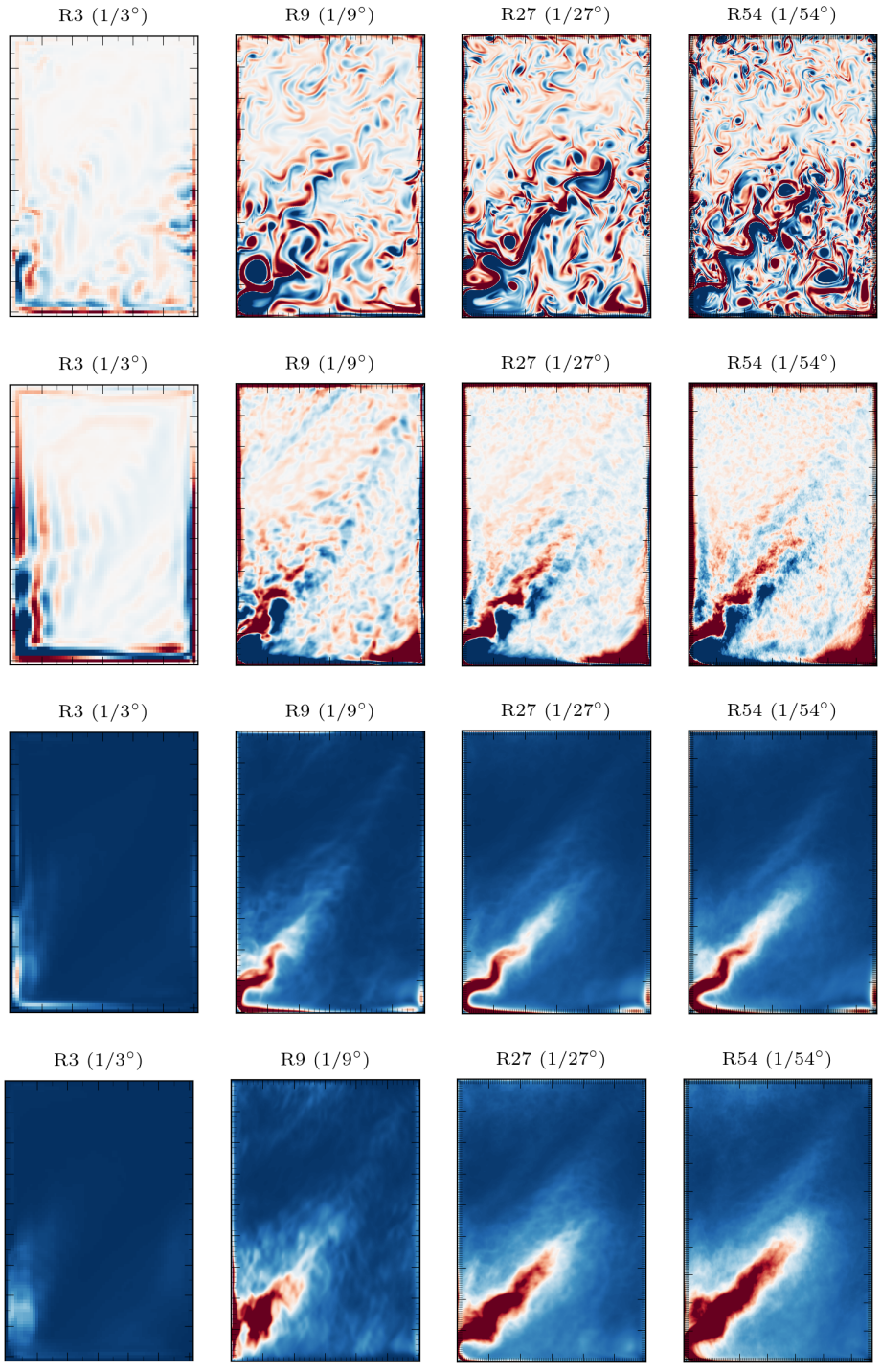


Figure 4.2 – Comparison of the four resolutions: from left to right 1/3°, 1/9°, 1/27° and 1/54°. From top to bottom: (a) snapshot of vorticity, (b) 10-years averaged relative vorticity $\overline{\zeta}^{10Y} = \overline{(\partial_x v - \partial_y u)/f}^{10Y}$, (c) 10-years averaged mean kinetic energy $\overline{\mathrm{MKE}}^{10Y}$ and (d) (c) 10-years averaged turbulent kinetic energy $\overline{\mathrm{TKE}}^{10Y}$ (with fluctuations computed with respect to the 10-years average).

4.1 Analysis of the deterministic experiments

Figure 4.2a shows a snapshot in time of the four different resolution deterministic simulations. The difference between the four cases is striking. The R3 simulation (first picture from the left) presents high activity only on the left border, where viscosity prevents the separation of the energetic structures from the boundary (Paiva et al., 2000). Proceeding to the right, the R9 simulation has a resolution that allows the formation of the jet current. However, the influence of the western boundary is still strong and the jet stream is found to interact with the boundary much more than in the high resolution simulations. Conversely, the two high resolution simulations, R27 and R54, present a well formed jet current detaching from the border at the same location (or with very small variation around this point). Concerning the smaller structures, high resolutions simulations such as R27 and R54 are rich of small scale eddies, fronts and filaments. At the beginning of eddying regime, such as the R9 simulation, these small structures are less prominent and their appearance is limited for brief periods of time. As expected, the coarse simulation R3 is completely incapable of producing these structures. These speculation about the time persistency of these structures can be assessed from Figure 4.2b, depicting the 10-years average relative vorticity $\overline{\zeta}^{10Y} = \overline{(\partial_x v - \partial_y u)/f}^{10Y}$. It shows primarily the difference between a high resolution simulation and a coarse resolution simulation. In the left panel, the R3 low resolution (35.3 km) simulation presents the aforementioned strong influence of the boundaries in the circulation patterns, while the interior of the domain has a low activity. Activity in the interior of the domain is present in all other regimes, represented by the R9, R27 and R54 simulations. However it is noticeable that the scales of eddies is larger in the R9 simulation, and becomes smaller and smaller with increasing resolutions. It is clear that, while R9 can sustain small eddies, these structures are not persistent for long enough times to be captured by the time average.

Figures 4.2c and 4.2d confirm what explained above, in terms of mean kinetic energy (c) and turbulent kinetic energy (d). The R3 simulation shows almost no activity in the interior of the domain in both figures, while highly energetic activity is present in the left border. Conversely, the R9 simulation shows that both mean and turbulent kinetic energy have increased and that the jet stream is present in the bottom left corner even though its behaviour is not stable, its meandering is strong and irregular and thus its average length is short. Moreover, close to the detachment area the average is faded and the turbulent kinetic energy is strong close to the western boundary, showing that its

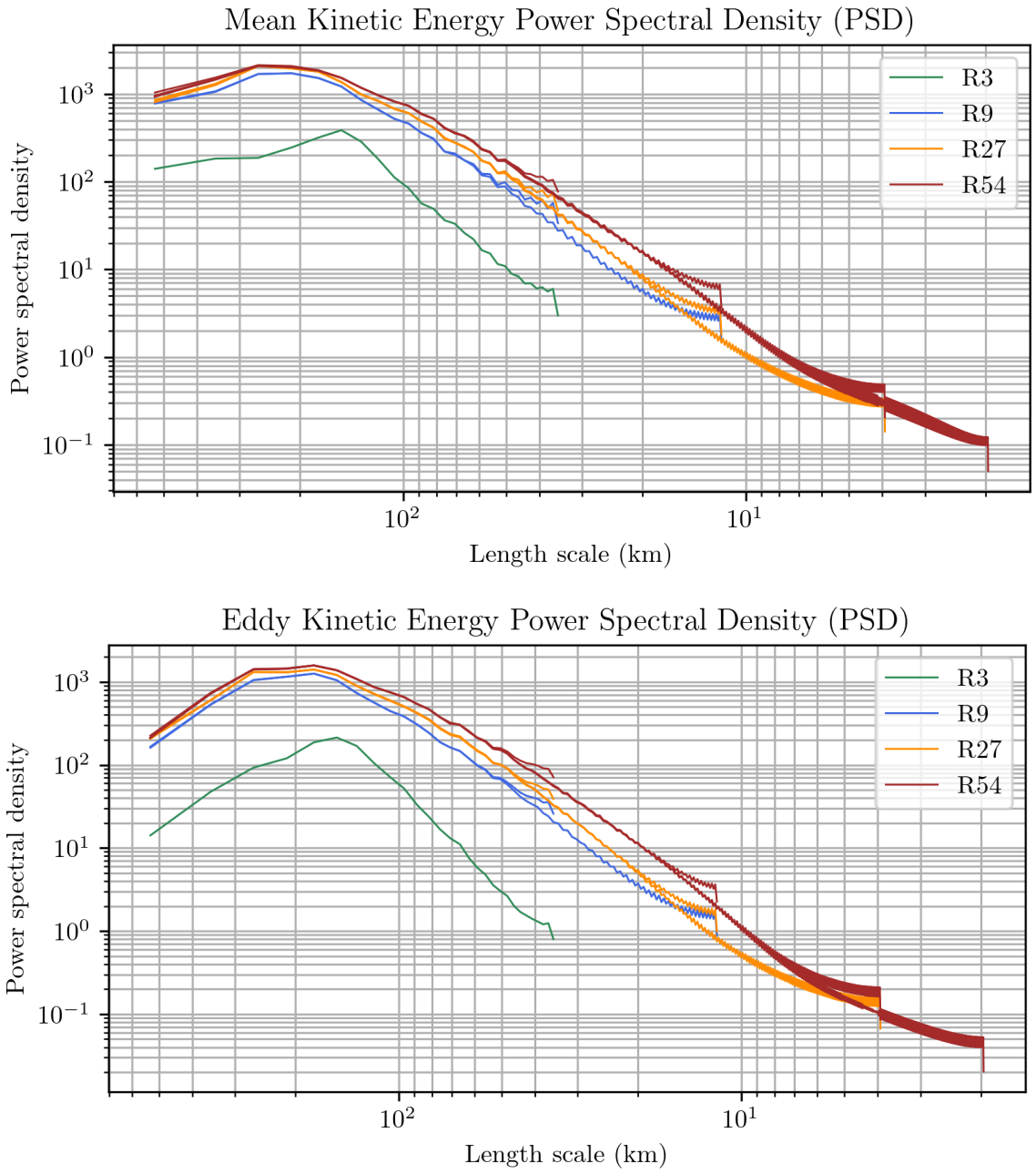


Figure 4.3 – Comparison of the 10-years averaged mean kinetic energy (top) and turbulent kinetic energy power spectral density (PSD). Each color represent a different resolution. The effects of grid coarsening is shown in the same color.

influence is not completely dissipated. In particular, a secondary jet is sometimes present in this configuration, not as stable as the primary, and well represented only at R27 and R54 resolutions (see the snapshots in Lévy, Resplandy, et al., 2012, for comparison), while the R9 shows interaction between the primary and secondary jet again through the boundary. Fine resolutions simulations such as R27 and R54 present instead a stable jet stream with average intensity well localized in space, signature of a long lasting and slowly evolving jet stream. Moreover, the turbulent kinetic energy is located around the jet stream , with a persistent turbulent activity near the secondary jet stream.

Figure 4.3 shows the mean (top) and turbulent (bottom) kinetic energy power spectral density. From this plot it is striking the difference between the low resolution model R3 and the rest of the simulations, as both the mean and turbulent kinetic energy spectrums are one order of magnitude less intense than the others, perfectly showing the reduced capabilities of this kind of model. Considering finer resolutions, it is systematic that the energy content of a simulation is smaller than that of a higher resolution simulation, even though the difference is not as large as when switching from a non-eddying regime such as R3 to an eddying regime such as R9 or more.

4.2 Stochastic modelling of the eddy-permitting resolution double gyre

The first target of the stochastic modelling is that of reproducing large scale features in coarse resolution models. As shown in the previous section the simulation lacks large scale features like the wind-driven jet stream. For this reason, the aim of the following experiments is that of enriching the behaviour of the flow with large scale characteristics hoping to enrich the fine scales as well. The choice of the noise model must thus be calibrated to reflect this application. On-line data models are unsuitable for this scope, as the simulation is not rich enough to self-sustain itself: off-line data models are thus to be preferred. Four experiments with different noises are presented. Two of these experiments are performed with a purely data-driven approach, namely the Proper Orthogonal Decomposition noise (from now on R3pod) defined as in Equation (3.41), the Dynamical Mode Decomposition noise (R3dmd) defined as in Equation (3.64). A third experiment is performed with a purely on-line data approach, the Vertical Profile Prescription noise (R3vpp) as defined in Equation (3.93). The last experiment merges the two approaches, with a dynamical mode decomposition based on sea surface elevation η to define the

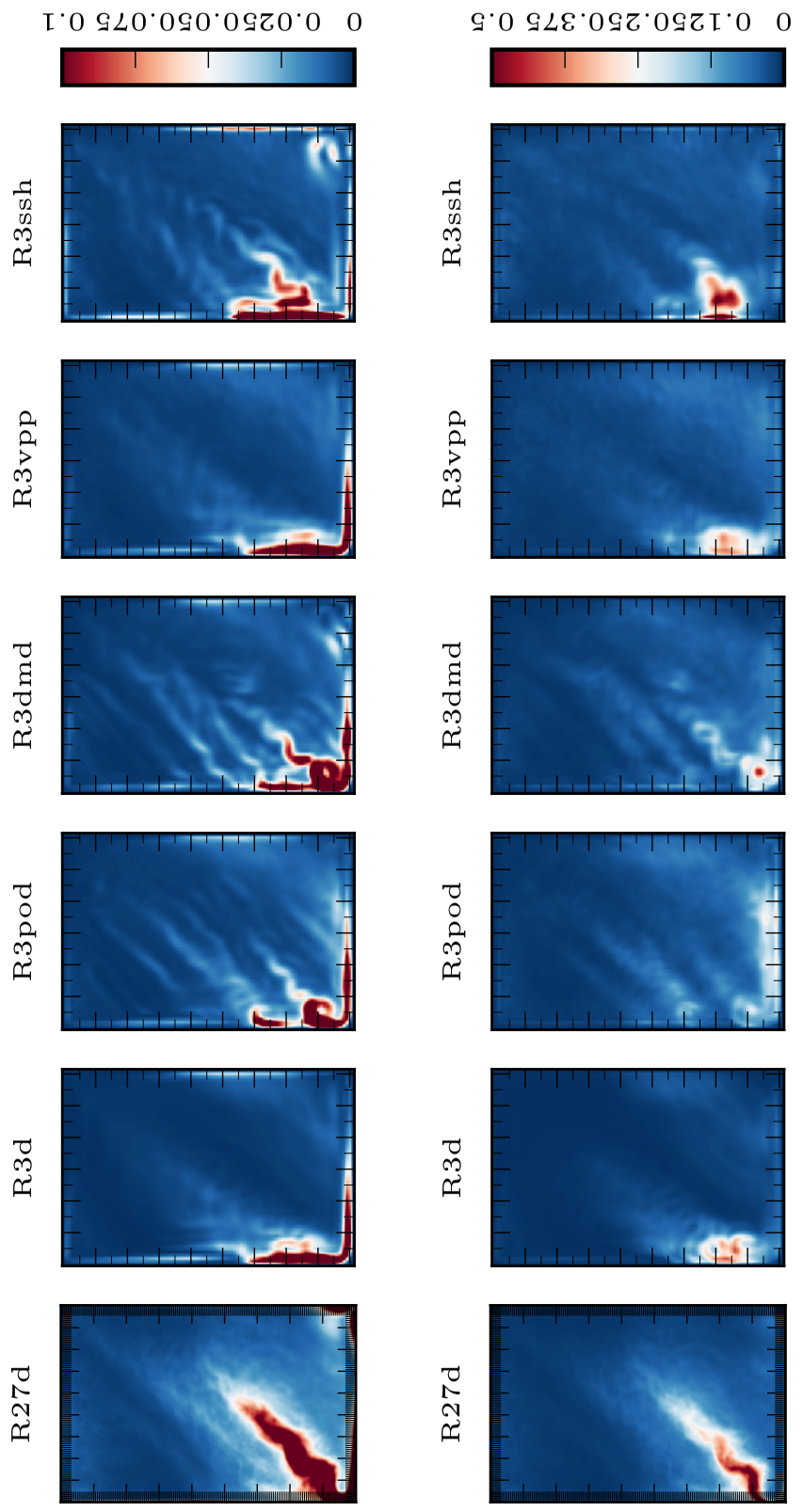


Figure 4.4 – Comparison of the 10-years averaged total kinetic energy $\overline{\mathrm{KE}}^{10Y}$ (top) and turbulent kinetic energy $\overline{\mathrm{TKE}}^{10Y}$ (bottom) for the different noise models employed at 1/3° resolution.

Girsanov drift, combined with a vertical profile noise. This last noise model (R3ssh) is constructed on the basis of the geostrophic velocity $\boldsymbol{u}_\eta$ reconstructed from the sea surface elevation η as

$$u_\eta = -\frac{g}{f}\frac{\partial \eta}{\partial y}, \qquad v_\eta = \frac{g}{f}\frac{\partial \eta}{\partial x}. \tag{4.1}$$

The dynamical mode decomposition can be run on this temporal field in order to capture the slow, correlated dynamics of η, so that a Girsanov drift can be defined and used to enrich of large scale behaviour while the previously described vertical profile noise can be used to enrich le small scale variability. In this case, the noise reads:

$$\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t = -\left[\overline{\mathbf{u}}_\eta^t + \sum_{m\in\mathcal{M}^c} \exp\left(\mathrm{i}\omega_m^\eta t\right) \boldsymbol{\varphi}_m^\eta\left(\boldsymbol{x}\right)\right] \mathrm{d}t + c\left(\boldsymbol{x}, z\right) \mathcal{G} * \mathrm{d}\mathbf{B}_t\left(\boldsymbol{x}\right). \tag{4.2}$$

The initial data η considered for this procedure is the synthetic data provided by the R27 simulation. To minimize the losses due to the interpolation of η on the grid of the velocity, the computation of the geostrophic velocity is performed on the R27 grid and subsequently filtered and downsampled to apply the DMD procedure. In real-life applications, the sea surface heigh observation data is commonly used in data assimilation procedures due to its availability with modern satellites. This noise constitutes thus an optimal candidate for further implementations in realistic configuration with access to real observations, not considered in this initial work. For all these noises, with only exception of R3vpp and the VPP component in R3ssh, both Girsanov and noise components are projected onto isopycnal surfaces.

4.2.1 Averaged total and turbulent kinetic energy

Qualitative analysis can be performed by discussing the 10-years temporal average of total and turbulent kinetic energy. For each performed simulation these maps are shown in Figure 4.4. The difference between the methods is striking, and one can conclude that for a task such as representing a large scale feature in coarse resolution simulations the off-line data driven methods are to be preferred over a purely on-line data driven method such as VPP. This latter method is beneficial to enhance the small scale turbulence in the interior domain, but it is not capable of self-aggregation to create larger structures, it only introduces more variability where the vertical velocity varies the most. Off-line data driven methods such as POD and DMD on the other hand show a remarkable imprinting of the data in the long term statistics of the simulation. A systematic bias in the positioning of

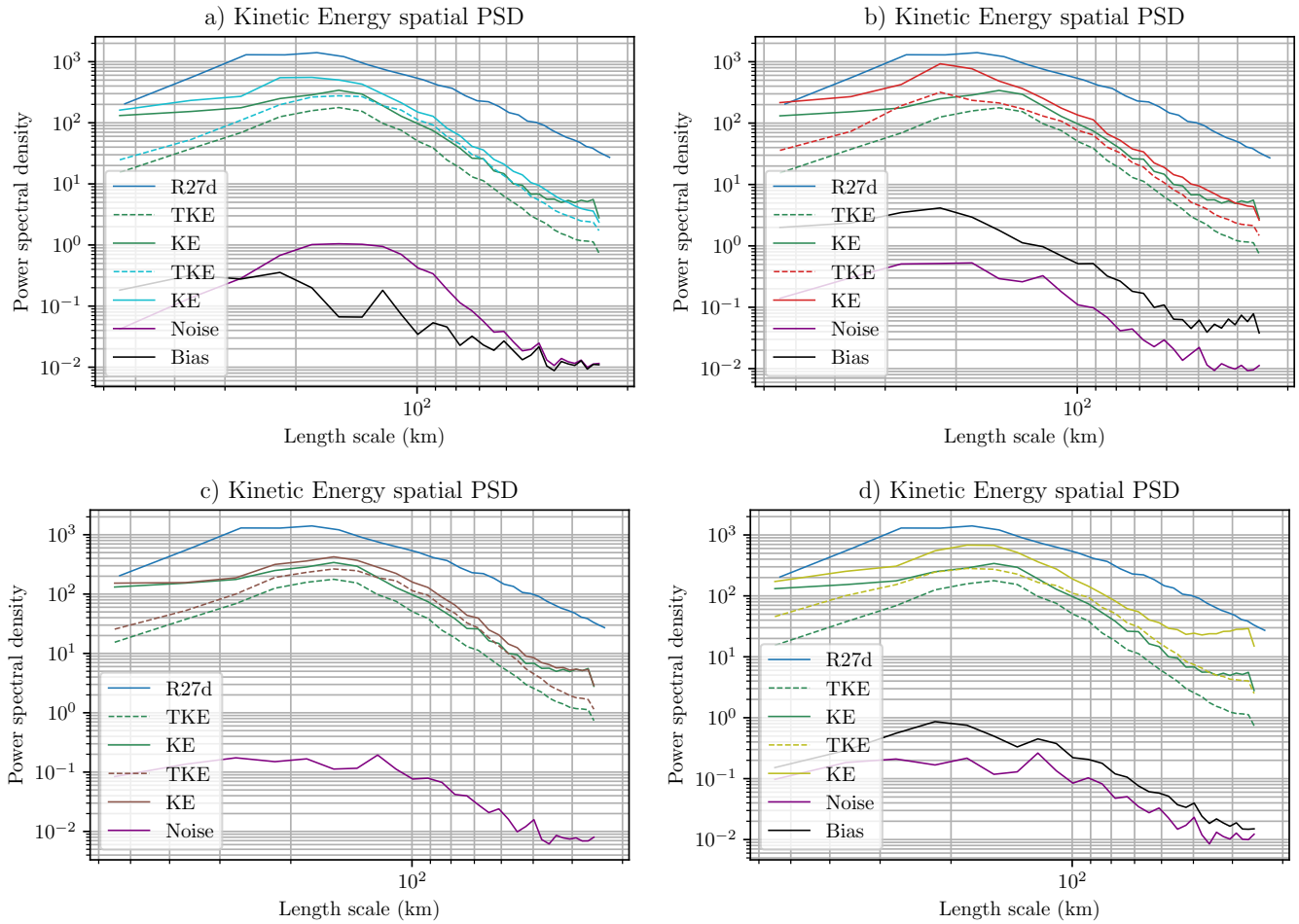


Figure 4.5 – Comparison of the 10-years averaged kinetic energy power spectral density for (a) R3pod, (b) R3dmd, (c) R3vpp and (d) R3ssh. The solid line shows the behaviour of total kinetic energy, the dashed line shows the turbulent kinetic energy. In all the plots, the green curves (solid and dashed) represent the reference R3 deterministic simulation. The solid blue line shows the reference R27 total kinetic energy spectrum.

the jet-stream detachment is visible: techniques to reduce this bias are being studied (see *e.g.* what discussed in Section 4.4).

4.2.2 Energy spatial power spectrum density

The power spectrum density for kinetic energy and turbulent kinetic energy are analysed. All cases presents substantial difference with respect to the reference. The R3pod case shown in Figure 4.5a shows that the method is beneficial in increasing both the global kinetic energy content and the turbulent kinetic energy content. The Girsanov drift, or bias, is in this case a time stationary field corresponding to the time average of the reference data, appropriately filtered and downsampled. Its energy content is mostly in the large scales, while at small scale only some residual is present. The noise field has its predominant energy content laying in the intermediate range of scales, possibly as a result of the bandpass filtering described in Section 3.3. A similar result is obtained with the DMD noise in the study case R3dmd. In this case, the Girsanov drift is no more stationary and it improves the large scale representation of the flow. The same benefits can be seen for the R3ssh model, where the bias term highly enhances the large scale energy content while the small scale, vertical profile noise increases the turbulent spectrum consistently at all scales. As shown in Figure 4.5c, both KE and TKE in this case are augmented.

The vertical profile noise R3vpp is shown in Figure 4.5d. It is noticeable that the turbulent kinetic energy is increased, while kinetic energy is not substantially improved. Moreover, this vertical profile noise introduces energy in almost the same amount for scales larger than 100 km, while sensibly dropping at scales below 100 km, due to the intrinsic nature of this noise as it is based on a Gaussian convolution filter spanning 3 grid cells of 35.3 km width. At this scale the shape of the spectrum differs from classical references and suggests that the additional energy that has been introduced does not act in synergy with the system but rather simply superimpose its effects through linear terms. The averaged fields in Figure 4.4 show that, if off-line data driven noises show good agreement with the targeted process, for the case of vertical profile noise there is no agglomeration of the small scales towards a larger coherent structure, hence the turbulence injected is directly dissipated. For this reason, a pure vertical profile noise is not envisageable at this low resolution, if not paired with some data driven approach or if not coupled with some external process. Finally, from the intercomparison of the different noise models, as shown in Figure 4.6, it can be established that the off-line data driven methods such as R3pod, R3dmd and R3ssh increase the energy content at large scale as

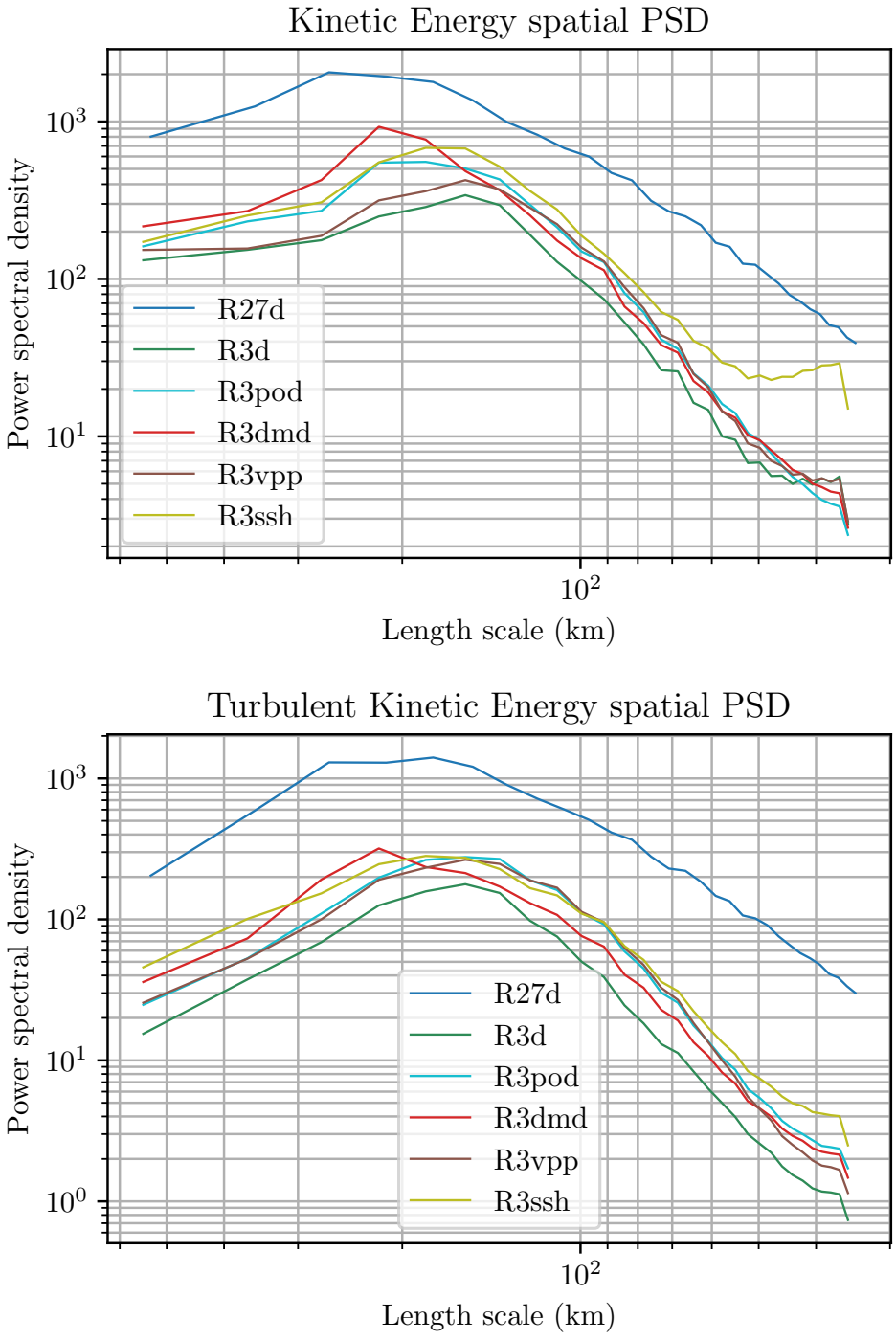


Figure 4.6 – Comparison of the 10-years averaged total kinetic energy $\overline{\mathrm{KE}}^{10\mathrm{Y}}$ (top) and turbulent kinetic energy $\overline{\mathrm{TKE}}^{10\mathrm{Y}}$ (bottom) power spectral density for the different noises

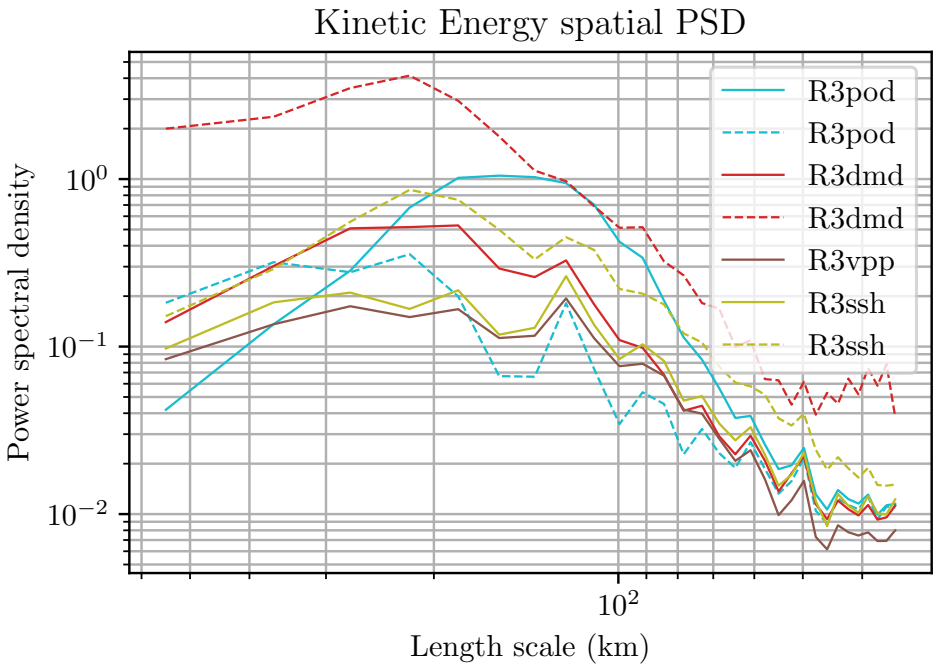


Figure 4.7 – Comparison of the 10-years averaged kinetic energy power spectral density for the different noises and corresponding Girsanov drifts.

intended, greatly enhancing both total and turbulent kinetic energy. At almost all scales and particularly over 100 km, the energy content is increased. As argued before, the purely data driven method of R3vpp simply superimposes its energy content to the spectrum. It is interesting to notice that similar behaviour of R3pod and R3vpp in the turbulent kinetic energy spectrum, suggesting the validity of POD for small scale turbulence. Figure 4.7 shows a comparison between the spectrums of the noise an bias term for each noise choice described so far. In particular, the two noise models relying on vertical profile noise follow the same construction and provide a very similar spectrum. R3pod and R3dmd noise terms have completely different structures, with the former targeting inertial scales and the latter having a profile closer to classical turbulent spectrums. This difference is most likely to be induced by the minimization problem defined in terms of energy of POD: it shows clearly that the POD is not extracting the largest scales of the flow but rather the most energetic scales that lie between 100 and 200 km. Concerning the Girsanov drifts, it is interesting to notice that R3ssh and R3dmd have different spectrums and that R3dmd introduced much more energy at large scales.

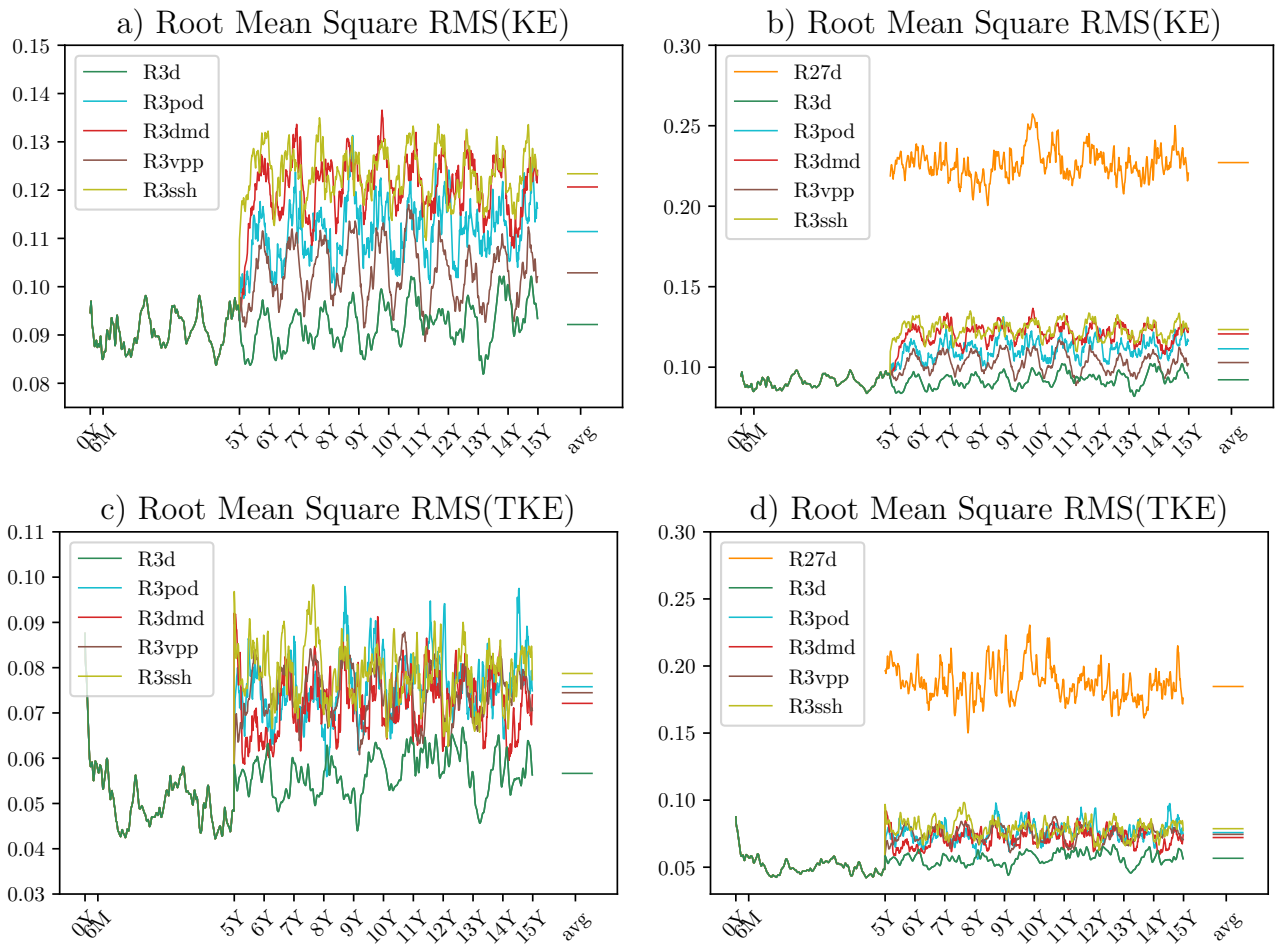


Figure 4.8 – Area-integrated total and turbulent kinetic energy in time, from top to bottom: (a) total kinetic energy fluctuations in time and (b) in perspective with the high resolution; (c) turbulent kinetic energy fluctuations in time and (d) in perspective with the high resolution.

Nevertheless, at small scales the R3dmd introduces too much energy, while R3ssh maintains a good slope. This behaviour can be noticed also in Figure 4.6 as indeed the kinetic energy of R3ssh is less dissipated at small scales. The difference in energy content is attributed to the distinct nature of the data: fully barotropic-baroclinic in the case of R3dmd and solely geostrophic-barotropic for R3ssh.

4.2.3 Energy vs time behaviour

In this section the behaviour in time of the integrated kinetic energy is considered. Without loss of generality, the surface layer is taken as representative of the full basin behaviour. The integrated kinetic energy in time is defined as

$$\mathrm{KE}\left(z=\eta,t\right)=\frac{1}{2}\frac{1}{\Omega}\int_{\Omega}u^{2}\left(\boldsymbol{x},z=\eta,t\right)+v^{2}\left(\boldsymbol{x},z=\eta,t\right)\,\mathrm{d}\boldsymbol{x}, \tag{4.3}$$

where the vertical component is not considered for simplicity. Consequently, the integrated turbulent kinetic energy in time is

$$\mathrm{TKE}\left(z=\eta,t\right)=\frac{1}{2}\frac{1}{\Omega}\int_{\Omega}u'^{2}\left(\boldsymbol{x},z=\eta,t\right)+v'^{2}\left(\boldsymbol{x},z=\eta,t\right)\,\mathrm{d}\boldsymbol{x}. \tag{4.4}$$

Figures 4.8a,b show the temporal behaviour of the integrated kinetic energy (4.3), while figures 4.8c,d show that of turbulent kinetic energy (4.4). The reference deterministic simulation R3det and the four stochastic simulations are run for a total of 15 years, where the stochastic components are switched on after 5 years of adjustment from the initial condition. For consistency, the fluctuations of the first 5 years are computed with respect of the average of the first 5 years only, while after the introduction of the stochastic parametrization the fluctuations are computed with respect to the average of the stochastic period alone, hence from year 5 to year 15. The energy of the high resolution R27 velocity data is shown in figures 4.8b,d, it corresponds to the first 10 years but it has been shifted for visualization purposes. A first trivial observation is that during the adjustment, the five low resolution models provide the same integrated KE and TKE, as expected. During this period, KE remains of the same magnitude, while TKE suffers a steep decay in the initial 6 months, due to the dissipation of the higher energy of the initial condition. After this highly dissipative period, TKE stabilizes around a much lower energy level than the initial condition. Considering the deterministic case R3det, both the levels of KE and TKE persist until the end of the simulation almost unvaried. This is not

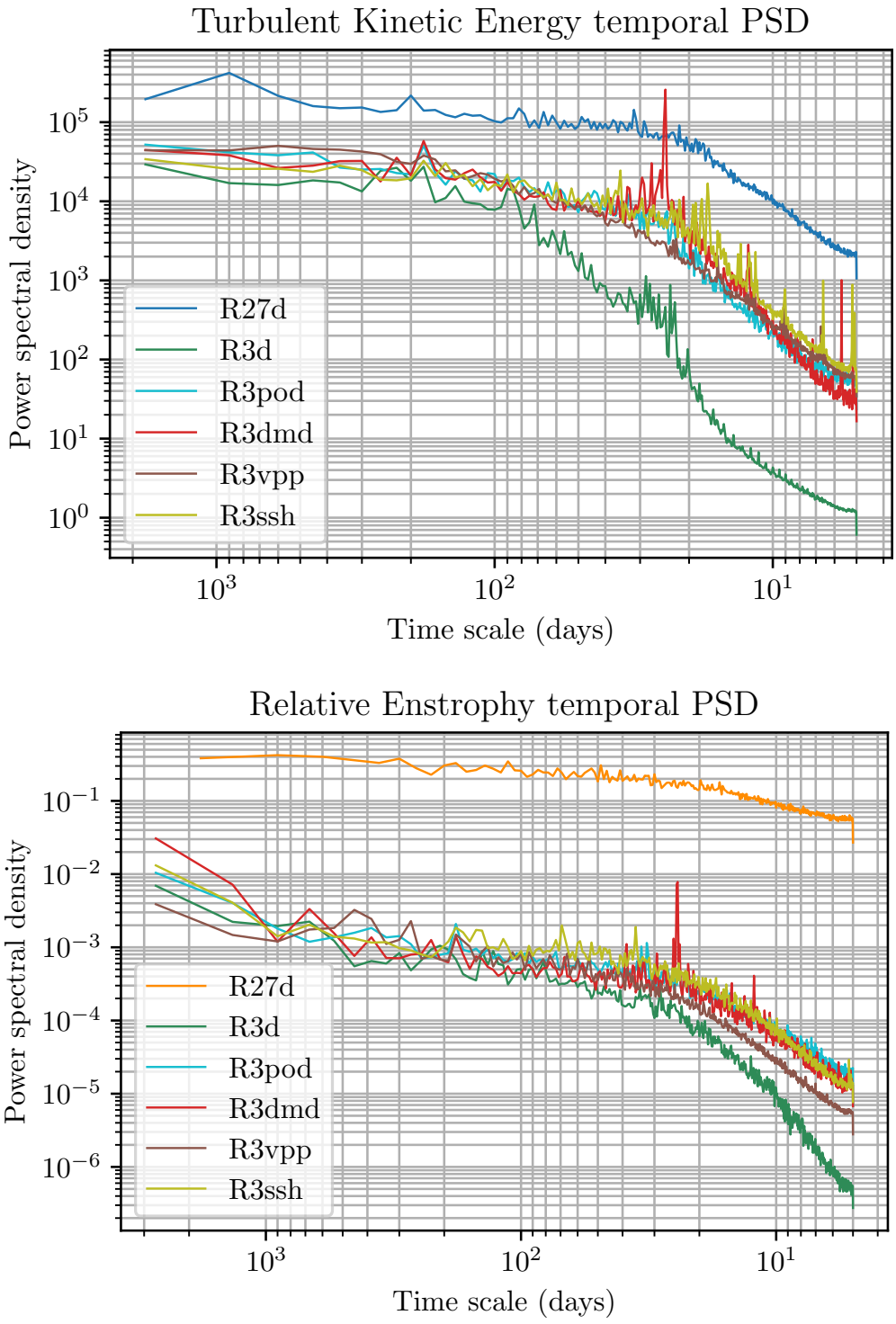


Figure 4.9 – Comparison of the 10-years averaged turbulent kinetic energy $\overline{\mathrm{TKE}}^{10Y}$ (top) and enstrophy $\overline{\zeta^2}^{10Y}$ temporal power spectral density for the different noises

the case of the four stochastic simulations. Around year 5 an increase of total and turbulent kinetic energy is present, although depending on the noise model chosen. Data driven models such as R3pod, R3dmd and R3ssh show a rapid increase of KE after the stochastic parametrization starts, while the R3vpp model performs better than deterministic but can be considered on below average for the KE. As expected, DMD-based methods strongly increase the kinetic energy. The TKE shows however a different behaviour, since the DMD data do not provide for fast fluctuating energy scales but rather slow components, thus R3dmd shows a low level of fluctuations and R3pod is the only data-based model that enhances fluctuations, advocating again for its use in fast stacle turbulence. The vertical profile noise model R3vpp enhances the simulation in fluctuations in a blind fashion and while enriching this statistics, it cannot be regarded as useful alone, in this context. Finally, the combined sea surface heigh DMD with vertical profile noise outperforms all the previously mentioned techniques. It is also interesting to notice that seasonality is present in the behaviour of all models, provided by the isopycnal projection and the vertical profile assignment. Finally, on the right of the plots the average level of energy is shown, with time average computed on the period from 5 to 15 years only, to leave aside the common deterministic initial period. As expected, R3dmd and R3ssh performs better that all other methods in terms of averaged KE, and it is significant that the improvement in averaged KE is comparable to that of R3dmd while using two-dimensional surface velocity data rather than three dimensional velocity data. R3pod is found to have a lower kinetic energy than these two simulations. R3gss enhances the kinetic energy on average, but the lack of scale aggregation shown by Figure 4.9 demonstrates that this enhancement is far from being useful. Concerning turbulent kinetic energy, as expected all methods are outperformed by R3ssh, that combines the benefits of a data driven model with that of a data-agnostic method.

4.2.4 Temporal power spectrum density

In Figure 4.9 the temporal spectra of the coarse resolution simulations are shown, in comparison to the fine resolution simulation. The first observable feature is that the low resolution deterministic simulation shows a shape that does not resemble the high resolution behaviour. At temporal scales smaller than 100 days the spectrum shows a steep decrease in kinetic energy intensity and a similar decay is shown for scales smaller than 50 days for the enstrophy, contrarily to what exhibited by the high resolution, that is a good energetic content at all scales above 20 days. This shows the low capability of

the coarse resolution simulation at resolving the ocean mesoscale. The second prominent feature is that R3vpp, the only on-line data based method, performs differently from all other methods: it increases the energy at the fastest scales, and gradually enhances slower scales. However, the lack of dynamics and self-aggregation capabilities jeopardise the effect of this increase of energy content. Conversely, off-line data driven methods show a greater effect in the mesoscale range, from 10 to 100 days. In particular, R3dmd shows peaks in the 20 to 40 days range, suggesting that the most energetic dmd modes are lying within this range. Moreover, it is interesting to notice such a sharp behaviour in this range, lies the time separation parameter τ_c presented in Section 3.6.2. R3ssh shows a remarkable resemblance with R3dmd, highlighting that simple sea surface elevation data can provide huge benefits at a reduced cost. The only energy based noise, R3pod, shows generalised good properties, especially at fast scales.

In conclusion, the off-line data based stochastic parametrization does not serve as a simple energetic offset towards a more rich state, but rather corrects the global shape of the energetic spectra. This property is not shared with the on-line data based parametrization, that at this resolution is not effective. This latter is just serving as a generalised energy and enstrophy content increase.

4.2.5 Metrics

From a quantitative point of view, the simulations are compared with different statistical metrics, computed for three model variables, namely the vorticity ζ, the horizontal kinetic energy $\frac{1}{2}\|\boldsymbol{u}\|^2$ and temperature T. Root Mean Square (RMS) is defined for the time average of the variable f_{M} of the noise model M as

$$\mathrm{RMS}(\overline{f}^{t}_{\mathrm{M}}) = \sqrt{\frac{1}{V}\int_V \overline{f}^{t\,2}_{\mathrm{M}}\,\mathrm{d}V}, \tag{4.5}$$

providing a measure of the energy content for the variable f_{M}. This metric first shows the difference in energy content of the high resolution and the low resolutions models. The improvement of the RMS can be evaluated using a simple linear function as

$$I(\mathcal{S}, \overline{f}^{t}_{\mathrm{M}}) = \frac{\mathcal{S}(\overline{f}^{t}_{\mathrm{M}}) - \mathcal{S}(\overline{f}^{t}_{\mathrm{R3d}})}{\mathcal{S}(\overline{f}^{t}_{\mathrm{R27d}}) - \mathcal{S}(\overline{f}^{t}_{\mathrm{R3d}})}, \tag{4.6}$$

Table 4.2 – Comparison of the values obtained with the diagnostic metrics: boldface values represent the best scenarios.

		ζ	$\frac{1}{2}(u^2+v^2)$	T
RMS($\overline{f}^t$)	R27d	9.27e-06	0.118	19.1
	R3det	1.63e-06	0.025	18.62
	R3pod	1.82e-06	0.031	**18.82**
	R3dmd	2.18e-06	0.040	**18.81**
	R3vpp	1.54e-06	0.025	18.68
	R3ssh	**2.30e-06**	**0.052**	18.70
I(RMS, $\overline{f}^t_{\text{M}}$)	R3pod	2%	6%	**41%**
	R3dmd	7%	16%	**39%**
	R3vpp	-1%	0%	12%
	R3ssh	**8%**	**29%**	17%
RMSE($\overline{f}^t$)	R3det	5.14e-06	0.101	**0.83**
	R3pod	**5.07e-06**	0.097	0.92
	R3dmd	5.18e-06	**0.089**	0.94
	R3vpp	5.09e-06	0.100	0.85
	R3ssh	5.12e-06	0.109	0.84
RMSE(σ)	R3det	8.07e-06	0.108	**0.32**
	R3pod	7.44e-06	0.105	0.38
	R3dmd	7.50e-06	**0.101**	0.39
	R3vpp	7.51e-06	0.107	**0.32**
	R3ssh	**7.31e-06**	0.106	0.33
PCC(f)	R3det	0.48	0.53	0.99
	R3pod	0.42	0.59	0.99
	R3dmd	**0.59**	**0.70**	0.99
	R3vpp	0.50	0.56	0.99
	R3ssh	0.49	0.29	0.99
GRE(f)	R3det	68.16	2225	0.27
	R3pod	10.69	91	**0.19**
	R3dmd	12.01	**76**	0.20
	R3vpp	14.94	325	0.25
	R3ssh	**9.37**	77	0.29

so that it takes value 1 when the stochastic simulation reaches values comparable to those of the R27 simulation or 0 when it does not perform better than the R3 deterministic simulation. A direct comparison between R27 and the stochastic and deterministic R3 models is provided by the Root Mean Square Error (RMSE). The time average of the high resolution data is downsampled without filtering on the low resolution grid (as it is supposed to be smooth enough), where it is compared to the low resolution time average as

$$\mathrm{RMSE}(\overline{f}^{\,t}_{\mathrm{M}}) = \sqrt{\frac{1}{V}\int_V \left(\left[\overline{f}^{\,t}_{\mathrm{R27}}\right]^{\downarrow}_{\mathrm{M}} - \overline{f}^{\,t}_{\mathrm{M}}\right)^2 \mathrm{d}V}. \tag{4.7}$$

Defining $\overline{f}^{\,t} = \frac{1}{T}\int_T f \,\mathrm{d}t$ the time average and $\sigma_f^2 = \frac{1}{T}\int_T (f - \overline{f}^{\,t})^2 \,\mathrm{d}t$ the time variance, the pattern correlation (PC, Grooms, A. Majda, et al., 2014)

$$\mathrm{PC}\,(f) = \frac{\displaystyle\int_V \sigma^2_{f,\mathrm{R27}}\sigma^2_{f,\mathrm{M}}\,\mathrm{d}V}{\left(\displaystyle\int_V \sigma^4_{f,\mathrm{R27}}\,\mathrm{d}V \int_V \sigma^4_{f,\mathrm{M}}\,\mathrm{d}V\right)^{1/2}} \tag{4.8}$$

and the Gaussian Relative Entropy (GRE, Grooms, A. Majda, et al., 2014) at a single point

$$\mathrm{GRE} = \frac{1}{2}\frac{1}{V}\int_V \left[\frac{(\overline{f}^{\,t}_{\mathrm{R27}} - \overline{f}^{\,t}_{\mathrm{M}})^2}{\sigma^2_{f,\mathrm{M}}} + \frac{\sigma^2_{f,\mathrm{R27}}}{\sigma^2_{f,\mathrm{M}}} - 1 - \ln\left(\frac{\sigma^2_{f,\mathrm{R27}}}{\sigma^2_{f,\mathrm{M}}}\right)\right] \mathrm{d}V, \tag{4.9}$$

can be computed, this latter measuring with a single criterion both the mean and variance reconstructions. The first term on the right-hand side of GRE represents the error in the mean weighted by the variance of the model. The remaining terms measure the error in model variability and is referred to as "dispersion". The lower this criterion the better the reconstruction. It can be observed from (4.9) that this criterion is minimal if, for all points, the mean is perfectly reconstructed and if the variance of the reference equals the one of the coarse model tested. Table 4.2 summarises the results obtained. Overall, these values confirm what has been presented so far. There is no metric that is sensibly improved for the on-line data based model R3vpp at this resolution, while R3dmd and R3ssh are the two models that perform the best. On average, both vorticity and kinetic energy contents are increased for every off-line data driven model, with the best performance provided by DMD-based models. If on the one hand the error in the time average might suffer from the incorrect positioning of the jet stream, as for the R3dmd, on the other hand different metrics to assess the benefit of the presence of the jet stream itself, such as heat transport evaluation, should be considered to better evaluate the validity of the method.

The same observation applies for the Patter Correlation PC, for which the incorrect jet stream positioning may exacerbate the differences with the high resolution model more than a deterministic model with no jet stream at all. Finally, Gaussian Relative Entropy, or GRE, balances the effects of mean and variance estimation, and show remarkable improvements for all stochastic parametrizations.

Conclusion

This chapter has illustrated the main benefits of LU for a coarse resolution simulation. In particular, all metrics chosen demonstrate a remarkable increase of the quality of the large scale features representation and furthermore an enhancement of the variability. Clearly, the choice of the noise model plays the most important role in the play: off-line data driven models are to be preferred to on-line data driven models, since the base state of the simulation does not contains enough energetic structures and so they must be induced from higher resolution simulations or other source of data, such as satellite observations. In this sense, using satellite altimeter data is prototyped in the R3ssh test case, where the sea surface elevation η provided by a higher resolution simulations constitutes a proxy for real altimetry data.

4.3 Stochastic modelling of the eddy-resolving resolution double gyre

In the eddy-resolving regime simulation, at a resolution of 11.8 km (1/9°) the targeted process is no more a large scale feature such as the wind driven jet current. This resolution already resolves this process with a grade of accuracy that, while leaving room for improvement, already allows for the targeting of smaller scales. The focus of these experiments is thus that of testing if the stochastic modelling at this resolution provides for a greater degree of variability than a deterministic simulation. For this reason, only on-line data based noises are chosen: the vertical profile prescription noise (R9vpp, see Section 3.9), the pseudo-observation POD (R9pso, see Section 3.7) and the wavelet-based noise (R9wlt, see Section 3.8). Figure 4.10 displays the behaviour of the three different stochastic models in terms of mean kinetic energy and turbulent kinetic energy averaged over a period of 5 years. The main features of the stochastic parametrization at this resolution are already present in this analysis: all the methods considered increase the kinetic

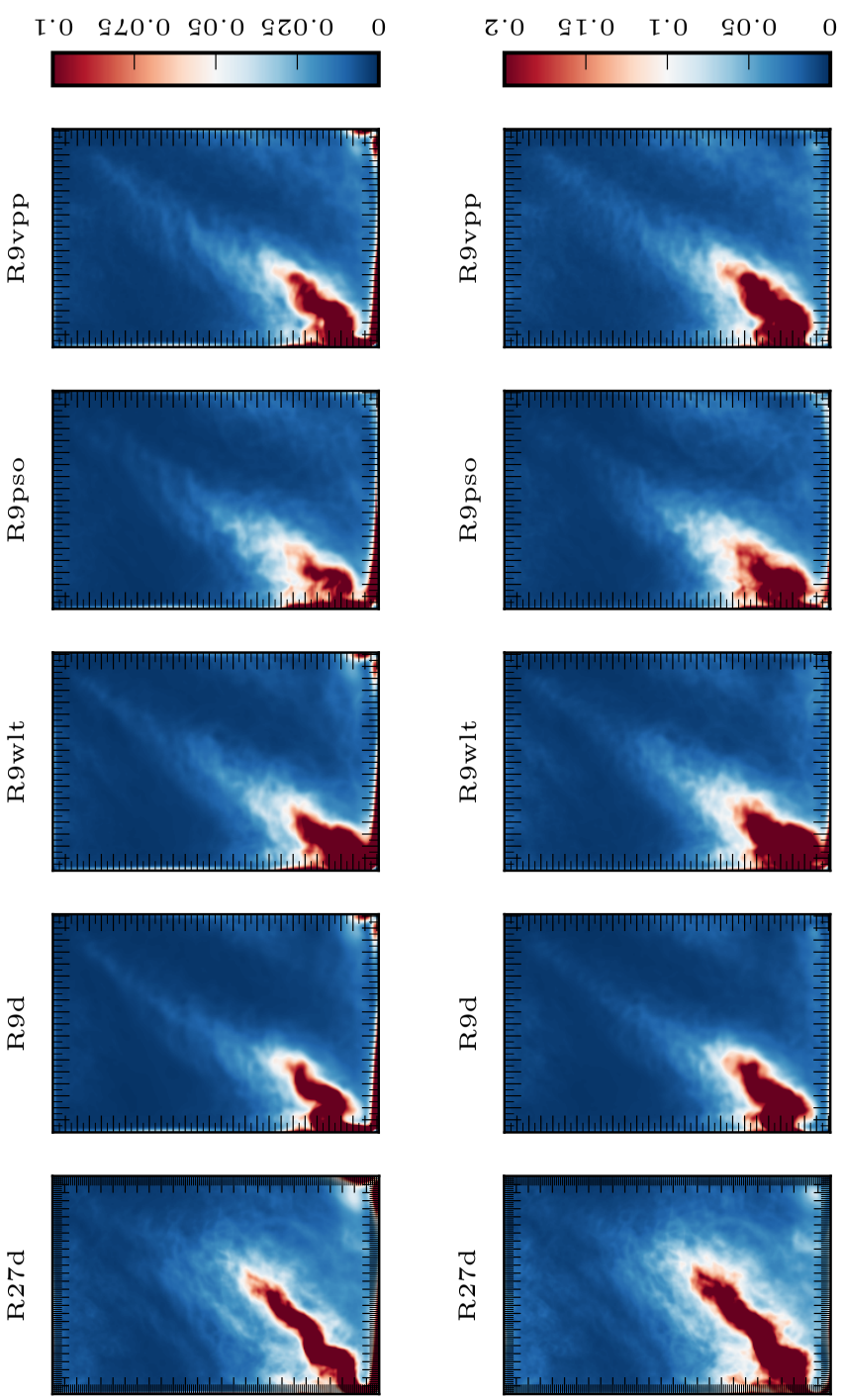


Figure 4.10 – Comparison of the 5-years averaged mean kinetic energy $\overline{\mathrm{MKE}}^{5Y}$ (top) and turbulent kinetic energy $\overline{\mathrm{TKE}}^{5Y}$ (bottom) for the different noise models employed at 1/9° resolution.

energy away from the jet stream in both mean and turbulent sense. However, R9wlt and R9pso show a tendency of destabilisation of the jet stream that corresponds to a reduction in terms of average length and, in the R9pso, in terms of intensity. Conversely, R9vpp outperforms the former two in both terms of average length and detachment position. The jet stream in this case is indeed interacting less with the western boundary, resulting in a more stable behaviour. This constitutes a key difference between the wavelet based method, the pseudo-observations method and the VPP method: the first two methods both rely on the current state velocity and they both extract features from this field. These procedure will create thus a noise well tied to the spatial structures of the resolved field, hence enhancing destabilisation around them. Conversely, VPP noise is tied to a proxy profile of the vertical kinetic energy, governing its amplitude, and it is more evenly applied over the domain. This drawback of the pseudo-observation noise model is well shown in Figure 4.12, where the energy content at large scale is not increased but rather diminished. A possible mitigation of this aspect might be the joint use of the structure following destabilization provided by pseudo-observation models and a more evenly spread vertical profile prescription models. From Figures 4.12 and 4.11 it is possible to assess that all the models work at small scales in both mean an turbulent kinetic energy. In this latter case, while at large scale the stochastic parametrisation only provides a slight increase in energy content, at smaller scale all the models are highly energetic, with the pseudo-observation model introducing much more turbulence at small scales than the other models. Such a conclusion holds as well in terms of relative enstrophy, *e.g.* Figure 4.11, where the drawback of R9pso is clearly visible. Concerning VPP noise, both Figures 4.11 and 4.12 show that its peak action is around scales of 30 km, which is approximately the width of the horizontal Gaussian filter employed.

4.3.1 Energy temporal power spectrum density

Figure 4.15 represents the temporal spectra for turbulent kinetic energy and relative enstrophy. The stochastic parametrisation provided by wavelet and VPP noise prove efficacy in enhancing both the energy and enstrophy content, with VPP outperforming the on-line velocity based noises at almost all scales, with an emphasis on the fastest scales. In agreement with the previous exposition, pseudo-observation noise is the least efficient amongst the novel parametrizations, with results comparable to those of the deterministic experiment.

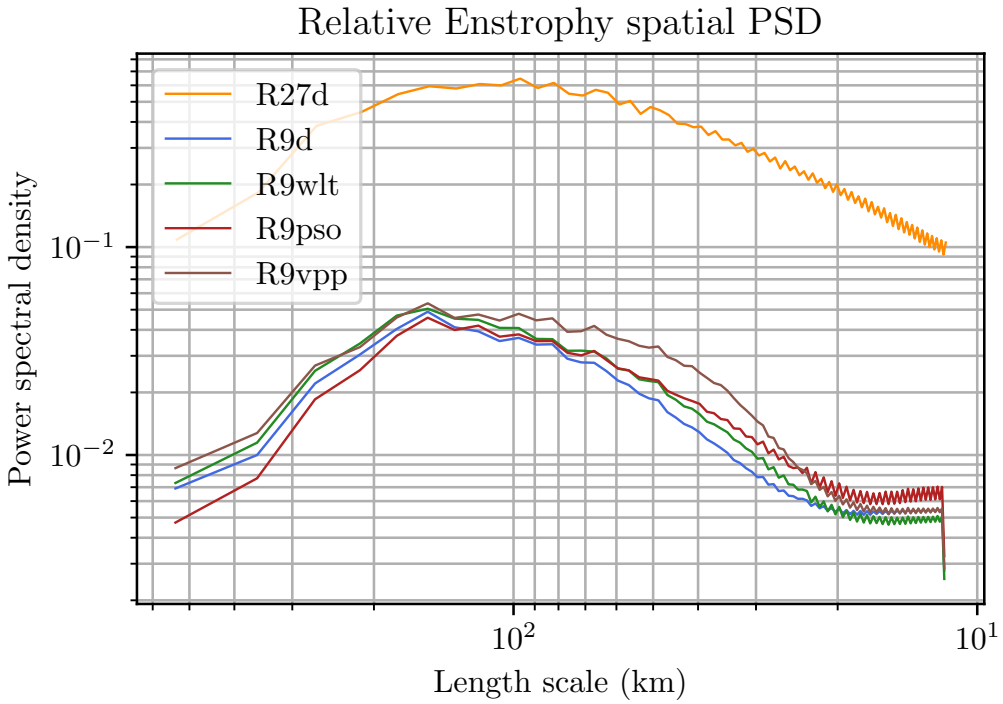


Figure 4.11 – Comparison of the 5-years averaged enstrophy spatial power spectral density for the noise models employed, with comparison to the 1/27° reference spectrum.

4.4 Stress testing: extreme noise

In this section a particular experiment involving the Vertical Profile Prescription noise is going to be described. As introduced in the previous sections, VPP noise possesses one striking difference with respect to the other on-line procedures: the amplitude of the noise is quasi-homogeneous in space. Others techniques presented, such as wavelet analysis or pseudo-observation POD are capable of localizing large energetic structures and thus the noise is stronger around these structures. While this property can be used actively to enhance the variance around high energy flows, it poses a challenge in the calibration of the noise term, since high amplitude is expected near large structures and a low amplitude is expected everywhere else. Moreover, computational bounds are implemented in the standard core of NEMO to stop the model when non-physical values are reached (these bounds are $S < 0$, $S > 100$psu, $|\boldsymbol{v}| > 10$m/s and $\eta > 20$m). These two aspects combined do not allow to arbitrarily increase the noise level for any given model: around highly energetic structures a high value of the noise will be attained and it might trigger the computational bounds, while the rest of the domain might not be affected by the noise.

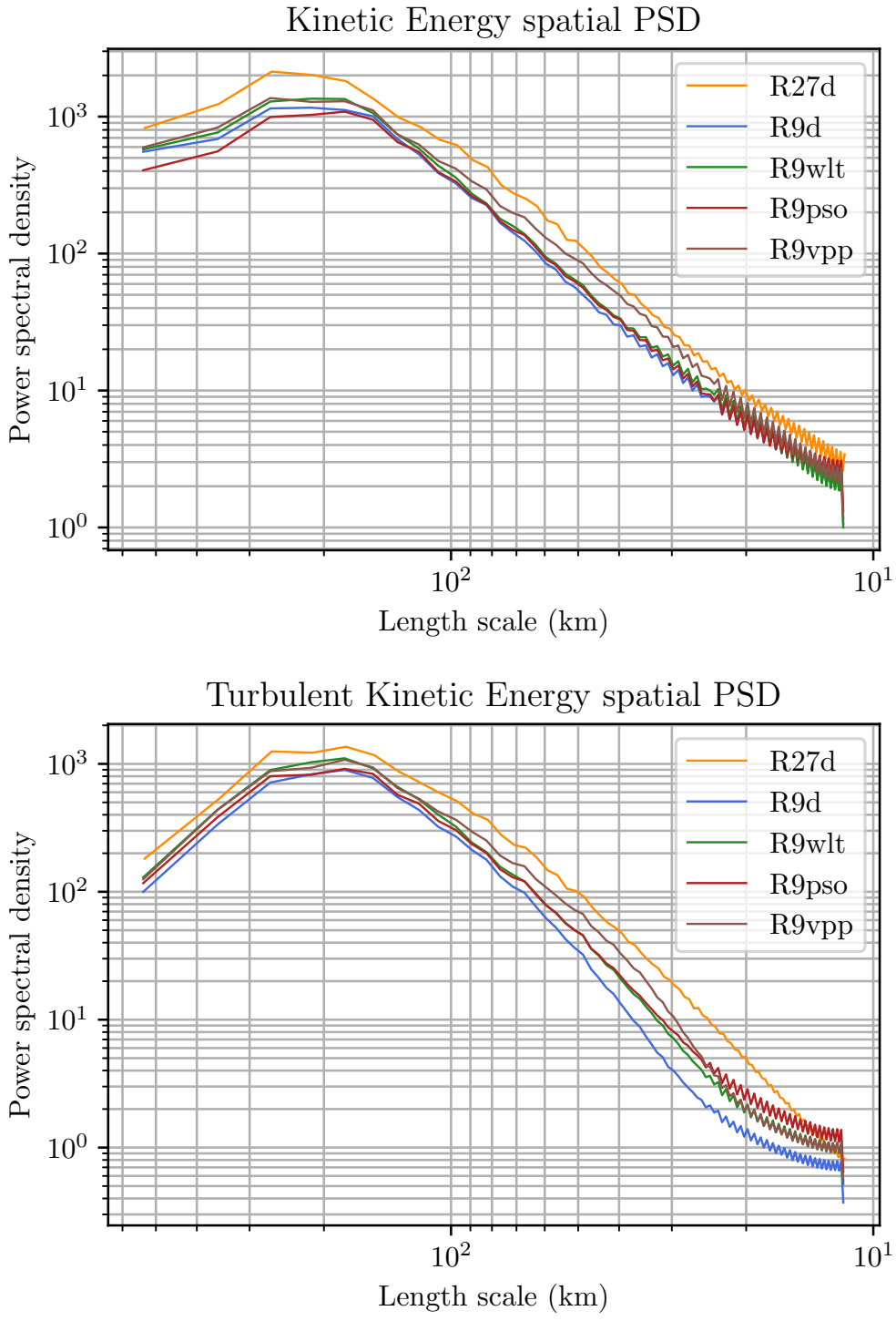


Figure 4.12 – Comparison of the 5-years averaged total (top) and turbulent (bottom) kinetic energy power spectral density for the noise models employed, with comparison to the 1/27° reference spectrum.

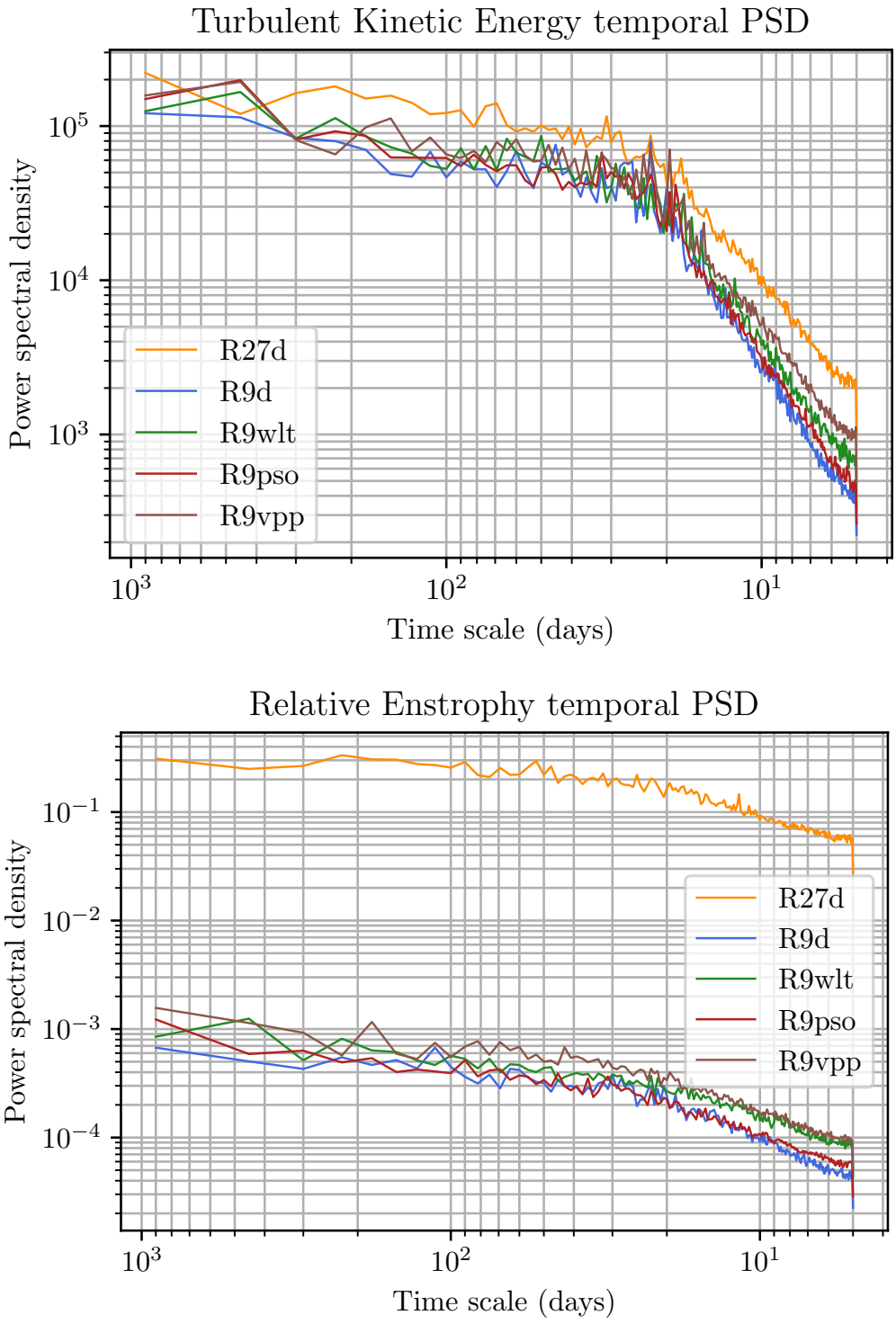


Figure 4.13 – Comparison of the 5-years averaged turbulent kinetic energy $\overline{\mathrm{TKE}}^{5Y}$ (top) and enstrophy $\overline{\zeta^2}^{5Y}$ temporal power spectral density for the different noises

Conversely, the VPP noise introduced in Section 3.9 is completely blind with respect to the horizontal structure of the mean flow, it only knows its vertical structure through the term $c(\mathbf{x}, z)$ of Equation 3.93. The consequence is that VPP noise is applied with the same structure and a slowly varying spatial amplitude over the whole domain. In other words, VPP noise can be set to high values without breaking the stability of the model. For this reason, it is interesting to perform simulations with a high amplitude VPP noise to test the model's stability and the model's reaction to such noise. It is paramount to stress that the amplitude of the noise here employed largely exceeds the amplitudes of the noises used in the previous sections. This is thus not intended as a physically meaningful test, but rather a computational stress test. First of all, from the snapshots analysis the large scale structures are found to be consistently reduced in their elongation and general coherence along the jet stream. However, large structures are not completely destroyed and a distinction between the northern and southern gyres is clear. If on the one hand this was quite expected, as the increase in small scale turbulence is known to be a trigger for an increase in dissipation, on the other hand the presence of the jet is a remarkable outcome. The spectrum shows an unrealistic response to the noise forcing and a high level of turbulence in both mean kinetic energy and eddy kinetic energy. All of these aspects are showing that the noise is poorly calibrated. Nevertheless, on a 1 year average the detachment of the jet stream is better located than the deterministic counterpart, with respect to the R27 detachment point. From a purely computational point, the simulation remains stable up to completion after 10 years, even though a trend of salinity decrease is visible. This trend can be estimated in less than 0.15 g/year and would cause the model to stop if the simulation is run long enough. This phenomenon is attributed to the strength of the noise and is an indicator of bad calibration. Summarising, despite providing non-physical results, this test shows that the method and its implementation work on a wide range of scenarios, from low to high noise amplitudes and from small to large scale noises and do not trigger computational instabilities that break the simulation. This strategy is however non advisable and the noise should be tuned to have a balanced effect. Moreover, VPP noise should not be used alone but in conjunction with other parametrization aiming at interacting with large scale features, such as a Girsanov drift or a current time velocity based noise model. Localising the VPP noise to some particular areas of the computational domain, such as the western boundary, potentially with an ad-hoc physical modelling choice for the vertical profile to be imposed, will be investigated as a strategy to help the large scale jet-stream to maintain a more regular path.

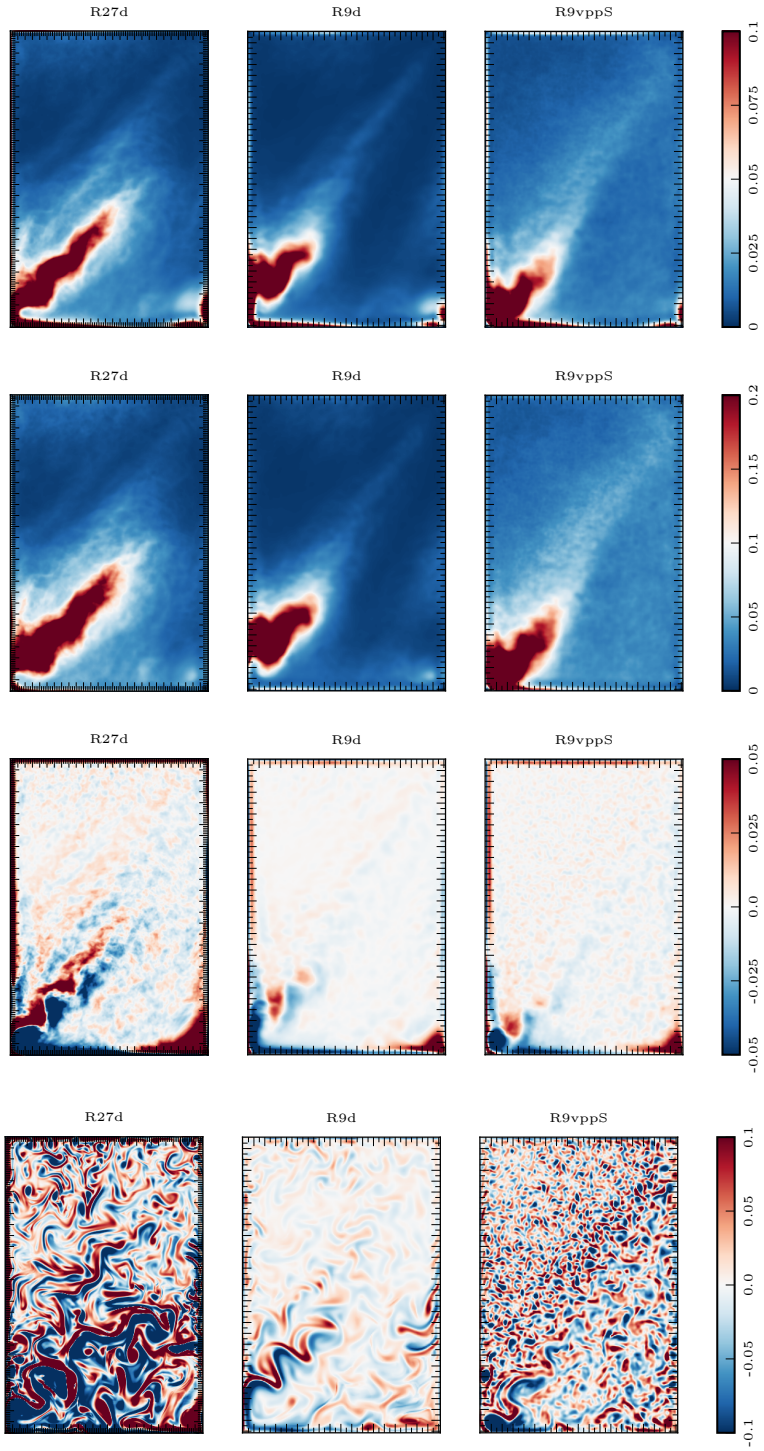


Figure 4.14 – Comparison of the 10 years averaged kinetic energy $\overline{\mathrm{KE}}^{10\mathrm{Y}}$ (first from the top), averaged turbulent kinetic energy $\overline{\mathrm{TKE}}^{10\mathrm{Y}}$ (second from the top), 10-years averaged relative vorticity $\overline{\zeta}^{10\mathrm{Y}} = \overline{(\partial_x v - \partial_y u)/f}^{10\mathrm{Y}}$ (third from top) and a snapshot in time of relative vorticity (bottom).

Conclusion

Looking at time averages of the vorticity it is possible to observe that the main problem of the R9 deterministic simulation is the interaction with the western boundary, caused by the strong and unstable meandering of the jet stream. This problem, at this specific resolution, is of very difficult solution. On-line data driven noises such as wavelet and pseudo-observation may alleviate this condition but they cannot solve it efficiently, as they intrinsically depend on large scale structures in their construction. Moreover, they suffer the same drawback of off-line data driven models for coarse resolution simulations: they condense high noise in particular areas of the domain while leaving most of the domain almost unaffected. If one the one hand this is extremely valuable to augment the variability around highly energetic structures, on the other hand the noise is tied to follow large structures and will not improve the simulation away from these. For this reason, a more data-agnostic method such as the vertical profile prescription noise might be beneficial coupled to on-line data driven methods, to balance a more widespread variability with a more process-targeted variability. This approach is not presented in this thesis but its experimentation is undergoing. Moreover, the usage of external data to introduce a restoring tendency towards the more steady position of the higher resolution jet-stream is undergoing consideration. This approach can follow that presented for coarse resolution simulations in the R3ssh case, introducing a restoring tendency towards the slow geostrophic dynamics through the Girsanov drift while introducing an enhanced variability through on-line data driven noises.

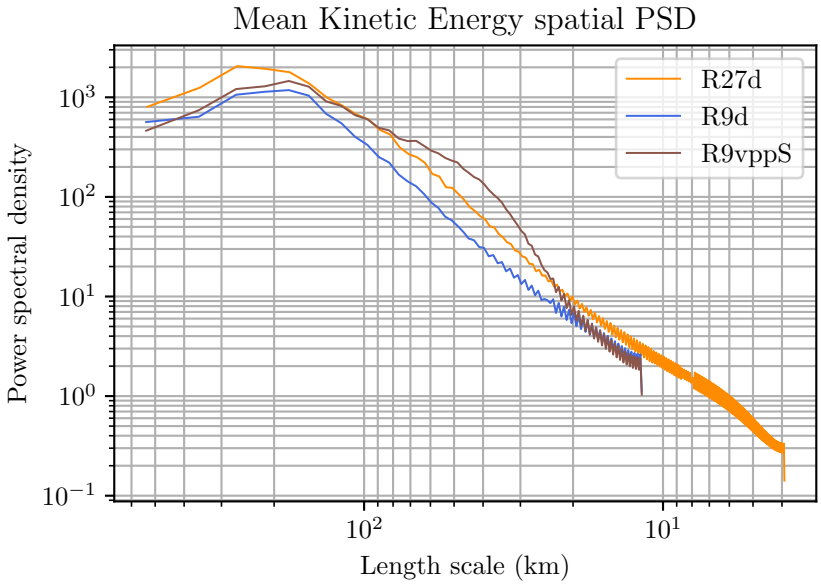


Figure 4.15 – Comparison of the 10-years averaged kinetic energy for the stress test case RvppS.

Conclusion

In this thesis, a stochastic representation of the small-scales of a fluid flow was discussed. This separation of scales, reminiscent in spirit of Reynolds splitting, constitutes the basis of the so-called Location Uncertainty framework: the Lagrangian displacement is split into a smooth-in-time component and a highly oscillating random term. From this ansatz, with extensive application of stochastic calculus, a stochastic partial differential equation (SPDE) is found to control the evolution in time of a process. In particular, a stochastic transport operator and a stochastic version of Reynolds Transport Theorem (RTT) can be introduced. The former can be identified as the equivalent of the material (or substantial) derivative of deterministic modelling in the novel stochastic framework: the underlying assumptions, those of describing the transport of a stochastic property along a path-line that is itself a stochastic process, correspond to the assumptions of the substantial derivative to describe the transport of a property while moving along a Lagrangian path-line. The differences in the two operators are induced by the effects of the newly introduced fast scales. These scales transport the stochastic property directly and indirectly, that is by means of statistically induced velocities as the Itô-Stokes drift, and have a fast-scale diffusion effect. From the principles established by the stochastic transport an extension to parcels of fluid can be found in the Stochastic Reynolds Transport Theorem (SRTT). This theorem establishes the fundamental behaviour of the temporal variation of a volume-integrated stochastic property when the volume itself is transported by a stochastic flow. This fundamental theorem provides means of interpreting the transport of measure, transport of stochastic quantities, conservation laws and balance equations in the stochastic framework. It is worth noticing that when the stochastic operators and theorems are applied to deterministic quantities, the description falls back to the classical deterministic description.

The stochastic transport principles established by SRTT were applied to describe the main equations describing a fluid flow and its fundamental processes, such as conservation of mass, conservation of momentum (intended as a balance between internal and external actions), conservation of energy and the generic conservation of an extensive property of the flow. With these ingredients, a stochastic Navier-Stokes equations can be derived.

Moreover, applying classical procedure such as asymptotic analysis, the Boussinesq system was introduced, allowing a simple description for flows characterised by a slowly varying density. In the deterministic framework this system constitutes the basis for most Ocean and Atmosphere dynamics modelling, and thus its stochastic development constitutes the grounding for the application of the LU principles to ocean and atmospheric prediction. Within this context, different simplification are applied in real-world studies, and this thesis aimed at gathering the most important models and their theoretical background. Amongst them, Primitive Equations (PE) and the Rotating Shallow Water system (RSW) are two key component of almost every Ocean or Atmosphere numerical core.

The implementation of the stochastic framework in the NEMO model was performed and tested in an idealised test case where the influence of the resolution is known to play a dominant role. The results obtained are promising. They first show that the mathematical framework of LU can be applied in its general terms to different problems with different objectives. This thesis considered two possible applications of LU to ocean modelling: that of completing large-scale simulation with highly energetic features that are normally not present in the state-of-the-art modelling and that of enriching a simulation of a greater degree of variability. These two tasks have different applications, as the former can be used in coupled Ocean-Atmosphere large scale models or long-term climate prediction, the latter might be exploited in short or medium term ensemble simulations to enrich the variability and increase the accuracy of the forecast. The choice of the noise model has proven to be critical, especially in relation with the scale of the simulation. The same noise might perform well at a given resolution and have little influence at another. For coarse resolution simulation, noise models that are *physics informed* are more suited to inform the simulation about unresolved scales from external data. The choice of the particular feature extraction procedure that is employed to generate the noise is open to be debated, as each technique has advantages and drawbacks. Nevertheless, this work show that relatively simple procedures such as POD or DMD can be applied to relatively small datasets, such as sea surface height, can have a performance comparable to that of the same techniques applied to larger datasets, such as baroclinic velocity. Moreover, noise models are not necessarily mutually exclusive, and optimal performance might be found balancing different techniques. At higher resolution, simple noises can increase the energy content and the variability. For these resolution a more thorough approach to noise modelling has to be established to understand better which processes to target.

Concluding, the LU framework introduces new degrees of freedom in the system that

can be exploited in different ways. The versatility of the method is such that its implementation does not require additional knowledge as compared to the maintenance or development of a general circulation model. Moreover, LU follows the specifics of the model in a natural way, such as following localised refinements. In addition, transport noise is a concept general enough to allow the description of the noise to be completely specified by the modeller. This freedom allows for targeting different objectives with ease within the same simulation, allowing potentially for different noise models complementing each other at different scales, different geographical position and tailored for different physical processes.

This thesis leaves open for multiple research paths in the future. The implementation of Location Uncertainty in different general circulation models will cast new light in the interaction between numerics and stochasticity. New techniques for feature extraction from external data can be pursued to efficiently inform the noise with physics in an inexpensive and efficient manner. Machine learning could be exploited as well inside the dynamical core to minimize the memory footprint of the noise model. From a less technical point of view, applications of the LU framework to less idealised configurations, such as the Mediterranean basin, are undergoing development.

Appendix A

Nucleus for European Modelling of the Ocean

Abstract

"Nucleus for European Modelling of the Ocean", NEMO, is a state-of-the-art modelling framework for ocean investigation and weather and climate forecasting, It is primarily developed and sustained by 5 European institutes: Centre national de la recherche scientifique (CNRS, France), Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC, Italy), Mercator Océan (France), Met Office (United Kingdom), National Oceanography Centre (NOC, United Kingdom). It is intended to be a flexible tool for studying the Earth climate system. In this chapter, the structure of the code is quickly reviewed and the main modifications implemented are detailed.

A.1 "Blue ocean" engine

NEMO, in its "blue ocean" dynamical core, solves the primitive equations of ocean thermodynamics. These equations set constraints on the dynamics of five prognostic variables, the three dimensional velocity field $\boldsymbol{v} = (\boldsymbol{u}, w)$, conservative temperature T and absolute salinity S. In the horizontal direction, the model is discretized on a curvilinear orthogonal grid and several options for the vertical grid are possible. The geographical coordinate system (λ, φ, z) used to introduce the primitive equations on a sphere in Chapter 1 has a singularity point at the North Pole that cannot be treated easily without requiring additional filtering. NEMO thus adopts a tripolar grid by merging the latitude-longitude

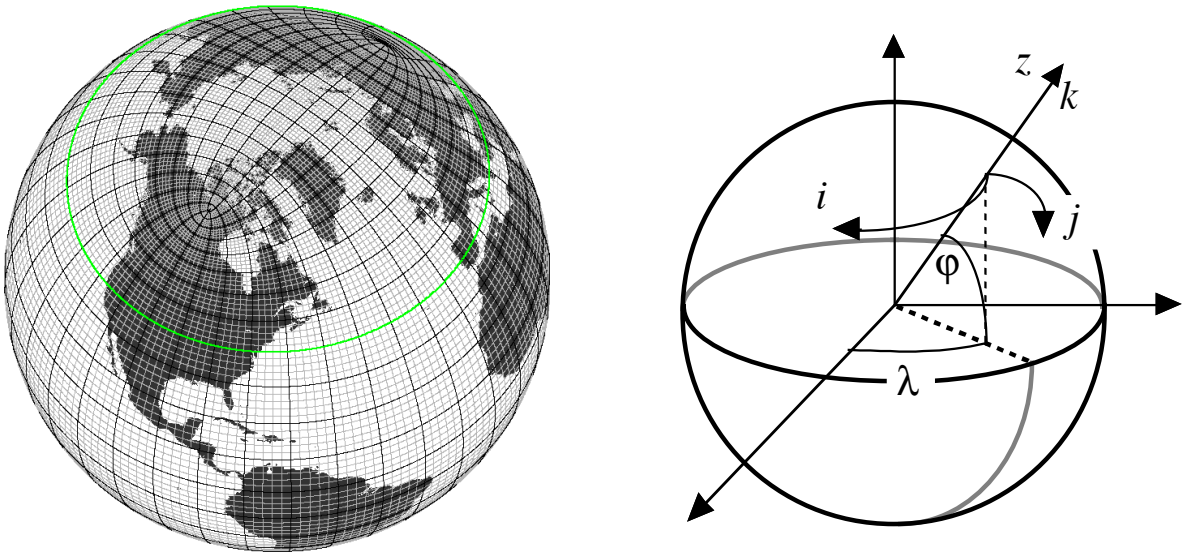


Figure A.1 – On the left, the tripolar grid introduced to solve for the polar singularity. On the right, the geographical coordinate system (λ, φ, z) as compared to the curvilinear coordinate system (i, j, k) implemented in NEMO. (Images taken from official NEMO documentation)

coordinates below 20°N with a custom, non-singular grid for higher latitudes. The standard ORCA grid built upon this procedure is shown in Figure A.1. As a consequence, an appropriate coordinate transformation must be introduced to handle in a uniform way the operations over any point of the grid. Let (λ, φ, z) be the geographical coordinate system in which a position is defined by the latitude $\varphi(i, j)$, the longitude $\lambda(i, j)$ and the distance from the centre of the Earth $a + z(k)$ where a is the Earth's radius and z the altitude above a reference sea level (A.1). The local deformation of the curvilinear coordinate system is given by e_1, e_2 and e_3, the three scale factors:

$$
\begin{aligned}
e_1 &= (a+z)\left[\left(\frac{\partial\lambda}{\partial i}\cos\varphi\right)^2 + \left(\frac{\partial\varphi}{\partial i}\right)^2\right]^{1/2}, \\
e_2 &= (a+z)\left[\left(\frac{\partial\lambda}{\partial j}\cos\varphi\right)^2 + \left(\frac{\partial\varphi}{\partial j}\right)^2\right]^{1/2}, \\
e_3 &= \left(\frac{\partial z}{\partial k}\right).
\end{aligned}
\tag{A.1}
$$

In the following, all partial derivatives and partial differential operators will be expressed in this model coordinate system using the two fundamental operators defined as:

$$\begin{aligned} \nabla q &= \frac{1}{e_1}\frac{\partial q}{\partial i}\,\mathbf{i} + \frac{1}{e_2}\frac{\partial q}{\partial j}\,\mathbf{j} + \frac{1}{e_3}\frac{\partial q}{\partial k}\,\mathbf{k}, \\ \nabla \cdot \boldsymbol{a} &= \frac{1}{e_1\, e_2}\left[\frac{\partial (e_2\, a_1)}{\partial i} + \frac{\partial (e_1\, a_2)}{\partial j}\right] + \frac{1}{e_3}\left[\frac{\partial a_3}{\partial k}\right]. \end{aligned} \tag{A.2}$$

The finite difference grid is defined in accordance to the Arakawa C-grid staggering, for which velocities u, v and w are located at the centre of the faces of a finite volume box, pressure and traces located at the centre of the box and finally vorticity at the box vertical edges. Considering a general variable q at adjacent points, the discrete differencing and averaging operators at the midpoint between them are:

$$\delta_i\left[q\right] = q(i+1/2) - q(i-1/2), \tag{A.3}$$

$$\overline{q}^{\,i+1} = \Big(q(i+1/2) + q(i-1/2)\Big)/2. \tag{A.4}$$

Similar operators are defined with respect to $i+1/2$, j, $j+1/2$, k, and $k+1/2$.

The goal of this study is that of embedding the Location Uncertainty formalism in the NEMO core in order to approach simulations in less idealised configurations than those performed up to this day. The complete implementation, based upon the discussion of the previous chapters, reads

Horizontal momentum:

$$\mathbb{D}_t \boldsymbol{u}_{\text{H}} + f\boldsymbol{k}\times(\boldsymbol{u}_{\text{H}}\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\mathbf{B}_t) = -\frac{1}{\rho_0}\nabla_{\text{H}}\left(p - \frac{\nu}{3}\nabla\cdot\boldsymbol{v}_s\right)\mathrm{d}t - \nabla_{\text{H}}\left(\mathrm{d}p_t^{\sigma}\right) + \boldsymbol{D}^{\boldsymbol{u}}\,\mathrm{d}t + \boldsymbol{F}^{\boldsymbol{u}}\,\mathrm{d}t \tag{A.5a}$$

Sea surface evolution:

$$\mathrm{d}_t\eta + \nabla\cdot\Big[(H+\eta)\left(\boldsymbol{U}_{\text{H}} - \boldsymbol{U}^s_{\text{H}} + \boldsymbol{\Sigma}\mathrm{d}\mathbf{B}^{\text{H}}_t\right)\Big] - \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\eta)\,\mathrm{d}t = \frac{\mathrm{emp}}{\rho_0}\mathrm{d}t \tag{A.5b}$$

Hydrostatic balance:

$$p'(x,y,z) = \frac{\nu}{3}\nabla\cdot\boldsymbol{v}_s\Big|_{\eta_b}^{z} + \int_{\eta_b}^{z} b + \boldsymbol{v}_s\cdot\nabla w + \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla w)\,\mathrm{d}\zeta, \tag{A.5c}$$

$$\mathrm{d}p_t^{\sigma}(x,y,z) = \int_{\eta_b}^{z}\boldsymbol{\sigma}\mathrm{d}\mathbf{B}_t\cdot\nabla w\,\mathrm{d}\zeta, \tag{A.5d}$$

Temperature and salinity:

$$\mathbb{D}_t T = D^T + F^T \tag{A.5e}$$

$$\mathbb{D}_t S = D^S + F^S \tag{A.5f}$$

Incompressibility:

$$\nabla \cdot [\boldsymbol{v} - \boldsymbol{v}_s] = 0, \qquad \nabla \cdot \boldsymbol{\sigma} \mathrm{d}\mathbf{B}_t = 0, \tag{A.5g}$$

Equation of state:

$$\rho = \rho\left(T, S, p\right). \tag{A.5h}$$

The stochastic transport operator is written in flux form as

$$\begin{aligned} \mathbb{D}_t \theta = \mathrm{d}_t \theta + \frac{1}{e_1 e_2 e_3}\left[\frac{\partial\left(e_2 e_3 u^\star \theta\right)}{\partial i} + \frac{\partial\left(e_1 e_3 v^\star \theta\right)}{\partial j}\right] + \frac{1}{e_3}\frac{\partial\left(w^\star \theta\right)}{\partial k} \\ -\frac{1}{2}\frac{1}{e_1 e_2 e_3}\left[\frac{\partial\left(e_2 e_3 F_1^\theta\right)}{\partial i} + \frac{\partial\left(e_1 e_3 F_2^\theta\right)}{\partial j}\right]\mathrm{d}t - \frac{1}{2}\frac{1}{e_3}\frac{\partial\left(F_3^\theta\right)}{\partial k}\mathrm{d}t \end{aligned} \tag{A.6}$$

where $\boldsymbol{v}^\star = (\boldsymbol{v} - \boldsymbol{v}_s)\ \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t$ and the stochastic diffusion flux can be written in general form as

$$F_i^\theta = a_{i1}\frac{1}{e_1}\frac{\partial \theta}{\partial i} + a_{i2}\frac{1}{e_2}\frac{\partial \theta}{\partial j} + a_{i3}\frac{1}{e_3}\frac{\partial \theta}{\partial k}. \tag{A.7}$$

Expressions of $D_u^{\boldsymbol{U}}$, $D_v^{\boldsymbol{U}}$, $F_u^{\boldsymbol{U}}$, $F_v^{\boldsymbol{U}}$, D^T, D^S, F^T, F^S will be depending on the parametrization of the sub-grid diffusion and of the forcing.

A.2 NEMO LU implementation

The work of introducing LU in NEMO can be generally split in three macro-areas: the noise (and derived fields) definition and construction, the tailored implementation of the new differential operators needed by the method and the adaptation of the existing code to work within the new framework. This section attempts ad describing the details of all three aspects of the implementation. The implementation of the Location Uncertainty framework in NEMO requires the introduction of new fields in the dynamical core. These fields are namely the noise $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t$, the Itô-Stokes drift $\boldsymbol{v}_s$, the variance tensor $\boldsymbol{a}$ and if required the Girsanov drift $\boldsymbol{\mu}_t$. In particular, the noise, Itô-Stokes drift and Girsanov drift are three dimensional fields (thus requiring three components, each one depending on the three spatial coordinates), while the variance tensor is a symmetric 3×3 tensor thus representable with six fields (each of which depends from the three spatial coordi-

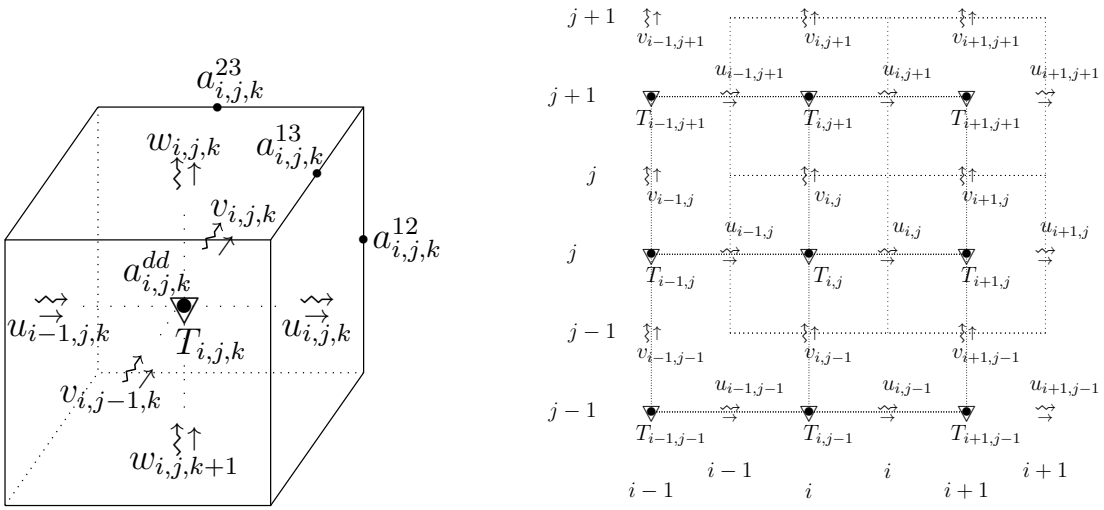


Figure A.2 – Schematic of the Arakawa C-Grid for the Location Uncertainty model. Left picture: at $T-$point, where the scalars are defined, the variance tensor $\boldsymbol{a}$ is also defined (and identified with the downward triangle). At u, v and $w-$points the corresponding noises are defined and identified with the squiggle arrow. Right picture: planar view of a discretization made with C-grid.

nates). The simplest case sees then the introduction of 15 new three dimensional arrays. When isopycnal projection is employed in conjuntion to a stationary assumption for the variance tensor, the memory cost is increased by 3 fields due to the necessity to store an initial Itô-Stokes drift that changes in time only with the application of the projection operator. If the Girsanov drift is subjected to a stationary assumption as well (such as in POD-based methods), then the same considerations are needed and the memory footprint is increased by 3 fields. An important preliminary aspect is the collocation of the new variables on the Arakawa C-grid. The positioning must be both consistent with the legacy code and be physically justified. The collocation of the noise $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t$ is the same as the velocity components of NEMO: $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{(x)}_t$ is located on the $u-$grid, $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{(y)}_t$ on the $y-$grid and $\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}^{(z)}_t$ on the $w-$grid. Things get more cumbersome when dealing with the variance tensor $\boldsymbol{a}$, where the choice is not well covered in ocean modelling literature. The choice made is to pose the diagonal components on the $T-$grid and the extra-diagonal components on the correspondent *mixed points*:

$$\boldsymbol{a} = \begin{bmatrix} a^{T}_{11} & a^{f}_{12} & a^{uw}_{13} \\ a^{f}_{21} & a^{T}_{22} & a^{vw}_{23} \\ a^{uw}_{31} & a^{vw}_{32} & a^{T}_{33} \end{bmatrix}. \tag{A.8}$$

This choice is motivated by the fact that the divergence of the stochastic diffusion is the Itô-Stokes drift, $\boldsymbol{v}^s = \frac{1}{2}\nabla \cdot \boldsymbol{a}$. When computing the discrete divergence of the diffusion tensor with this choice, the result is a velocity field naturally lying on the correct grid. Moreover, the sequence of operations needed to perform a double divergence leads to a quantity located on the $T-$grid. This is consistent with the definition of the compressibility effects

$$\nabla_{\text{H}} \left(\frac{\nu}{6} \nabla \cdot \nabla \cdot \boldsymbol{a} \right), \tag{A.9}$$

since if the quantity inside the parenthesis is located on the $T-$grid, then its gradient is correctly located on the u and $v-$grid. This choice is common in other research fields, such as porous media mechanics and seismology (*cfr.* Carcione, 1999). Equations (3.3) are chosen as templates for the noise construction, with a set of spatially depending modes $\phi^{(x)}$, $\phi^{(y)}$ and $\phi^{(z)}$ used to construct all related fields. Given that the the expansion is truncated at a level N, this adds $3N$ fields to the global memory footprint. With these considerations one can build each component of the noise ansatz efficiently as

$$\begin{aligned}
\boldsymbol{\sigma}_t (x,y,z)\, \mathrm{d}\mathbf{B}_t^{(x)} &= \sum_{n=1}^{N} \phi_n^{(x)} (x,y,z)\, \xi_n (t)\,, \\
\boldsymbol{\sigma}_t (x,y,z)\, \mathrm{d}\mathbf{B}_t^{(y)} &= \sum_{n=1}^{N} \phi_n^{(y)} (x,y,z)\, \xi_n (t)\,, \\
\boldsymbol{\sigma}_t (x,y,z)\, \mathrm{d}\mathbf{B}_t^{(z)} &= \sum_{n=1}^{N} \phi_n^{(z)} (x,y,z)\, \xi_n (t)\,,
\end{aligned}$$

while the variance tensor requires interpolation of the original velocity fields to be located on the prescribed locations:

$$\begin{aligned}
a^{xx}(x,y,z) &= \sum_{n=1}^{M} \overline{\phi_n^{(x)}(x,y,z)\,\phi_n^{(x)}(x,y,z)}^{\,i}, & a^{xy}(x,y,z) &= \sum_{n=1}^{M} \overline{\phi_n^{(x)}(x,y,z)}^{\,j}\, \overline{\phi_n^{(y)}(x,y,z)}^{\,i} \\
a^{yy}(x,y,z) &= \sum_{n=1}^{M} \overline{\phi_n^{(y)}(x,y,z)\,\phi_n^{(y)}(x,y,z)}^{\,j}, & a^{xz}(x,y,z) &= \sum_{n=1}^{M} \overline{\phi_n^{(x)}(x,y,z)}^{\,k}\, \overline{\phi_n^{(z)}(x,y,z)}^{\,i} \\
a^{zz}(x,y,z) &= \sum_{n=1}^{M} \overline{\phi_n^{(z)}(x,y,z)\,\phi_n^{(z)}(x,y,z)}^{\,k}, & a^{yz}(x,y,z) &= \sum_{n=1}^{M} \overline{\phi_n^{(y)}(x,y,z)}^{\,k}\, \overline{\phi_n^{(z)}(x,y,z)}^{\,j}.
\end{aligned}$$

The Itô-Stokes drift $\boldsymbol{u}_s$ can be computed seamlessly by applying the divergence operator by column

$$
\begin{aligned}
u_s &= \frac{1}{e_1 e_2 e_3}\left\{\delta_{i+1/2}\left[e_2 e_3 a_{11}\right] + \delta_j\left[e_1 e_3 a_{21}\right]\right\} + \frac{1}{e_3}\delta_k\left[a_{31}\right], \\
v_s &= \frac{1}{e_1 e_2 e_3}\left\{\delta_i\left[e_2 e_3 a_{12}\right] + \delta_{j+1/2}\left[e_1 e_3 a_{22}\right]\right\} + \frac{1}{e_3}\delta_k\left[a_{32}\right], \\
w_s &= \frac{1}{e_1 e_2 e_3}\left\{\delta_i\left[e_2 e_3 a_{13}\right] + \delta_j\left[e_1 e_3 a_{23}\right]\right\} + \frac{1}{e_3}\delta_{k+1/2}\left[a_{33}\right].
\end{aligned}
$$

Each of the noise models described in Chapter 3 can be described as a summation of modes $\phi_n(\boldsymbol{x})$, appropriately multiplied by a temporal coefficient $\xi_n(t)$. The global memory footprint is thus $15 + 3N$ new fields. It is clear that using methods that allow for a better representation with a small N (such as DMD) is fundamental and that operating with these fields becomes more cumbersome increasing the resolution of the model (thus justifying the usage of on-line data models). From the pure implementation point of view, a module containing the fields definition and their initialization is introduced to interface the legacy code of NEMO in its initialization. The operations thus far described, that are common to all methods, are contained in a second module that is called in the time stepping routine, where conditional constructs switch on or off the stochastic routines to compute and apply the noise fields. Each one of the methods exposed in chapter 3 is self-contained in a dedicated module that construct the noise fields. The original routines of NEMO that are affected are thus `nemogcm.F90` for the initialization (uses 3 new modules and conditionally calls 3 new routines), `step_oce.F90` for the inclusion of the necessary modules (7 new modules) and `step.F90` for the calls of the stochastic routines (1 call to noise generation, 4 conditional calls to the noise application and 1 diagnostic routine call). In the following, the new routines that has been implemented from scratch are described.

A.3 Stochastic advection discretization

Among all the procedures, horizontal advection is one of the most fundamental in the current implementation of the LU framework within NEMO. The advection of the horizontal velocity reads

$$
\mathcal{A}_{\boldsymbol{u}} = -\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla \boldsymbol{u}_{\mathrm{H}} = -\left[\frac{1}{e_1}\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{(x)}\frac{\partial \boldsymbol{u}_{\mathrm{H}}}{\partial x} + \frac{1}{e_2}\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{(y)}\frac{\partial \boldsymbol{u}_{\mathrm{H}}}{\partial y} + \frac{1}{e_3}\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t^{(z)}\frac{\partial \boldsymbol{u}_{\mathrm{H}}}{\partial z}\right] \quad \text{(A.10)}
$$

For simplicity of notation, and due to the generality of the procedure to be described, the advective velocity will be denoted with $\boldsymbol{u}_{\text{adv}}$ and the advected velocity will be denoted with $\boldsymbol{u}^{\text{in}}$. Due to the configuration of the standard Arakawa C-grid, the procedure needs a sequence of interpolation as the noise velocity and the gradient of the horizontal velocity are not located at the same point. The correct operational definition for the horizontal velocity is thus

$$\mathcal{A}_u = -\left[\overline{\frac{1}{e_{1t}}\overline{u_{\text{adv}}}^{\,i+1/2}\delta_{i+1/2}[u^{\text{in}}]}^{\,i} + \overline{\frac{1}{e_{2f}}\overline{v_{\text{adv}}}^{\,i+1/2}\delta_{j+1/2}[u^{\text{in}}]}^{\,j} + \overline{\frac{1}{e_{3uw}}\overline{w_{\text{adv}}}^{\,i+1/2}\delta_{k+1/2}[u^{\text{in}}]}^{\,k}\right],$$
$$\mathcal{A}_v = -\left[\overline{\frac{1}{e_{1f}}\overline{u_{\text{adv}}}^{\,j+1/2}\delta_{i+1/2}[v^{\text{in}}]}^{\,i} + \overline{\frac{1}{e_{2t}}\overline{v_{\text{adv}}}^{\,j+1/2}\delta_{j+1/2}[v^{\text{in}}]}^{\,j} + \overline{\frac{1}{e_{3vw}}\overline{w_{\text{adv}}}^{\,i+1/2}\delta_{k+1/2}[v^{\text{in}}]}^{\,k}\right].$$

The analogous procedure for tracers T, S is

$$\mathcal{A}_{T,S} = -\left[\overline{\frac{1}{e_{1f}}u_{\text{adv}}\delta_{i+1/2}[T^{\text{in}}]}^{\,i} + \overline{\frac{1}{e_{2t}}v_{\text{adv}}\delta_{j+1/2}[T^{\text{in}}]}^{\,j} + \overline{\frac{1}{e_{3vw}}w_{\text{adv}}\delta_{k+1/2}[T^{\text{in}}]}^{\,k}\right].$$

The aforementioned implementation of the advection procedure is employed when the equations of motion are written in energy-vorticity form. Whenever the flux form is preferred (as in the current version of the code), the original procedure to advect velocity and tracer is (duplicated and) modified in order to receive the advective velocity as an input field. This allows to compute the modified advection as

$$\boldsymbol{v}^{\star} = \boldsymbol{v} - \boldsymbol{v}^{s} + \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \tag{A.11}$$

and to provide it to the original routines to perform the operations.

A.4 Stochastic diffusion discretization

The second important operation to be performed is the diffusion of velocity and tracer due to the noise. This diffusion, being anisotropic, is not implemented in NEMO and so it has to be coded from scratch. Given a diffusive flux F, the diffusion is implemented by means of Gauss divergence theorem.

$$\int_V \nabla \cdot \boldsymbol{F}\,\mathrm{d}V = \int_{\partial V} \boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}s \quad \rightarrow \quad \int_V \nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}V = \int_{\partial V}(\boldsymbol{a}\nabla\theta)\cdot\boldsymbol{n}\,\mathrm{d}s \tag{A.12}$$

The positioning of the fluxes on the Arakawa C-grid is chosen in such a way that the divergence operation provides the result correctly located on the target grid. With this idea in mind, the fluxes of stochastic diffusion for velocity are

$$
\begin{aligned}
F_{u,x} &= a_{11}^{T} \partial_x u \, e_{2t} e_{3t}^{n} + \overline{\overline{a_{12}^{f} \partial_y u \, e_{2f} e_{3f}^{n}}}^{i,j} + \overline{\overline{a_{13}^{uw} \partial_z u \, e_{2u} e_{3uw}^{n}}}^{i,k}, \\
F_{u,y} &= a_{21}^{f} \overline{\partial_x u \, e_{1t} e_{3t}^{n}}^{i,j} + \overline{\overline{a_{22}^{T}}}^{i,j} \partial_y u \, e_{1f} e_{3f}^{n} + \overline{\overline{a_{23}^{vw}}}^{i,k} \overline{\partial_z u \, e_{1u} e_{3uw}^{n}}^{j,k}, \\
F_{u,z} &= a_{31}^{uw} \overline{\partial_x u \, e_{1t} e_{2t}}^{i,k} + \overline{\overline{a_{32}^{vw}}}^{i,j} \overline{\partial_y u \, e_{1f} e_{2f}}^{j,k},
\end{aligned}
$$

$$
\begin{aligned}
F_{v,x} &= \overline{\overline{a_{11}^{T}}}^{i,j} \partial_x v \, e_{2f} e_{3f}^{n} + a_{12}^{f} \overline{\partial_y v \, e_{2t} e_{3t}^{n}}^{i,j} + \overline{\overline{a_{13}^{uw}}}^{j,k} \overline{\partial_z v \, e_{2v} e_{3vw}^{n}}^{i,k}, \\
F_{v,y} &= \overline{\overline{a_{21}^{f} \partial_x v \, e_{1f} e_{3f}^{n}}}^{i,j} + a_{22}^{T} \partial_y v \, e_{1t} e_{3t}^{n} + \overline{\overline{a_{23}^{vw} \partial_z v \, e_{1v} e_{3vw}^{n}}}^{j,k}, \\
F_{v,z} &= \overline{\overline{a_{31}^{uw}}}^{i,j} \overline{\partial_x v \, e_{1f} e_{2f}}^{i,k} + a_{32}^{vw} \overline{\partial_y v \, e_{1t} e_{2t}}^{j,k},
\end{aligned}
$$

$$
\begin{aligned}
F_{w,x} &= \overline{\overline{a_{11}^{T}}}^{i,k} \partial_x w \, e_{2u} e_{3u}^{n} + \overline{\overline{a_{12}^{f}}}^{j,k} \overline{\partial_y w \, e_{2v} e_{3v}^{n}}^{i,j} + a_{13}^{uw} \overline{\partial_z w \, e_{2t} e_{3w}^{n}}^{i,k}, \\
F_{w,y} &= \overline{\overline{a_{21}^{f}}}^{i,k} \overline{\partial_x w \, e_{1u} e_{3u}^{n}}^{i,j} + \overline{\overline{a_{22}^{T}}}^{j,k} \partial_y w \, e_{1v} e_{3v}^{n} + a_{23}^{vw} \overline{\partial_z w \, e_{1u} e_{3uw}^{n}}^{j,k}, \\
F_{w,z} &= \overline{a_{31}^{uw} \partial_x w \, e_{1t} e_{2t}}^{i,k} + \overline{a_{32}^{vw} \partial_y w \, e_{1f} e_{2f}}^{j,k} + a_{33}^{T} \partial_z w \, e_{1T} e_{2T},
\end{aligned}
$$

while the fluxes of stochastic tracer diffusion are

$$
\begin{aligned}
F_{T,x} &= \overline{a_{11}^{T}}^{i} \partial_x T \, e_{2u} e_{3u}^{n} + \overline{a_{12}^{f} \overline{\partial_y T \, e_{2v} e_{3v}^{n}}^{i}}^{j} + \overline{a_{13}^{uw} \overline{\partial_z T \, e_{2t} e_{3w}^{n}}^{i}}^{k}, \\
F_{T,y} &= \overline{a_{21}^{f} \overline{\partial_x T \, e_{1u} e_{3u}^{n}}^{j}}^{i} + \overline{a_{22}^{T}}^{j} \partial_y T \, e_{1v} e_{3v}^{n} + \overline{a_{23}^{vw} \overline{\partial_z T \, e_{1u} e_{3uw}^{n}}^{j}}^{k}, \\
F_{T,z} &= \overline{a_{31}^{uw} \overline{\partial_x T \, e_{1t} e_{2t}}^{k}}^{i} + \overline{a_{32}^{vw} \overline{\partial_y T \, e_{1f} e_{2f}}^{k}}^{j}.
\end{aligned}
$$

The first implementation of Location uncertainty made use of this discretization. The current version leverages what introduced in Fiorini et al. (2022) and Boulvard et al. (2023), it approximates the diffusion with a double application of the advection (in non-conservative form)

$$
\frac{1}{2} \nabla \cdot (\boldsymbol{a} \nabla \theta) \sim \frac{1}{2} \boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla \left(\boldsymbol{\sigma}_t \mathrm{d}\mathbf{B}_t \cdot \nabla \theta \right). \tag{A.13}
$$

A.5 Projection on isopycnals

The isopycnal projector is defined as

$$\mathrm{P} = \delta_{ij} - \frac{(\partial_i \rho)(\partial_j \rho)}{\|\nabla \rho\|^2} = \frac{1}{\|\nabla \rho\|^2} \begin{bmatrix} (\partial_y \rho)^2 + (\partial_z \rho)^2 & -(\partial_x \rho)(\partial_y \rho) & -(\partial_x \rho)(\partial_z \rho) \\ -(\partial_x \rho)(\partial_y \rho) & (\partial_x \rho)^2 + (\partial_z \rho)^2 & -(\partial_y \rho)(\partial_z \rho) \\ -(\partial_x \rho)(\partial_z \rho) & -(\partial_y \rho)(\partial_z \rho) & (\partial_x \rho)^2 + (\partial_y \rho)^2 \end{bmatrix}, \tag{A.14}$$

with the aid of the small slope assumption, that means that the horizontal gradients are much smaller than vertical $\sqrt{(\partial_x \rho)^2 + (\partial_y \rho)^2} \ll \partial_z \rho$, the operator P becomes

$$\mathrm{P} = \begin{bmatrix} 1 & 0 & -r_1 \\ 0 & 1 & -r_2 \\ -r_1 & -r_2 & (r_1^2 + r_2^2) \end{bmatrix}, \tag{A.15}$$

with $r_1 = \partial_x \rho / \partial_z \rho$ and $r_2 = \partial_y \rho / \partial_z \rho$. These components are already available in NEMO. The projection $\mathrm{P}\boldsymbol{u}$ of the velocity reads:

$$\mathrm{P}\boldsymbol{u} = \begin{bmatrix} 1 & 0 & -r_1 \\ 0 & 1 & -r_2 \\ -r_1 & -r_2 & (r_1^2 + r_2^2) \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} u - \overline{\overline{r_1 w}}^{ik} \\ v - \overline{\overline{r_2 w}}^{jk} \\ -\overline{\overline{r_1 u}}^{ik} - \overline{\overline{r_2 v}}^{jk} + (r_1^2 + r_2^2)\, w \end{bmatrix}. \tag{A.16}$$

The projection $\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}}$ of the symmetric variance tensor $\boldsymbol{a}$ is composed of:

$$\begin{aligned}
(\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}})_{11} =& a_{11} + a_{33}\overline{r_1^2}^{k} - 2\,\overline{\overline{a_{13}}^{ik}\overline{r_1}^{k}} \\
(\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}})_{12} =& a_{12} - \overline{\overline{a_{23}}^{ik}\overline{r_1}^{j}} - \overline{\overline{a_{13}}^{jk}\overline{r_2}^{i}} + \overline{\overline{a_{33}}^{ij}\overline{r_1}^{j}\overline{r_2}^{i}} \\
(\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}})_{13} =& a_{13}\overline{(2r_1^2 + r_2^2)}^{i} + \overline{\overline{a_{23}}^{ij}\overline{r_1 r_2}^{i}} - \overline{\overline{a_{12}}^{jk}\overline{r_2}^{i}} - \overline{\overline{a_{11}}^{ik}\overline{r_1}^{i}} - \overline{\overline{a_{33}}^{ik}\overline{r_1\,(r_1^2 + r_2^2)}^{i}} \\
(\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}})_{22} =& a_{22} + a_{33}\overline{r_2^2}^{k} - 2\,\overline{\overline{a_{23}}^{jk}\overline{r_2}^{k}} \\
(\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}})_{23} =& a_{23}\overline{(r_1^2 + 2r_2^2)}^{j} + \overline{\overline{a_{13}}^{ij}\overline{r_1 r_2}^{j}} - \overline{\overline{a_{12}}^{ik}\overline{r_1}^{j}} - \overline{\overline{a_{22}}^{jk}\overline{r_2}^{j}} - \overline{\overline{a_{33}}^{jk}\overline{r_1\,(r_1^2 + r_2^2)}^{j}} \\
(\mathrm{P}\boldsymbol{a}\mathrm{P}^{\mathrm{T}})_{33} =& a_{33}\overline{(r_1^2 + r_2^2)^2}^{k} - 2\,\overline{\overline{a_{13}}^{ik}\,\overline{r_1\,(r_1^2 + r_2^2)}^{k}} - 2\,\overline{\overline{a_{23}}^{jk}\,\overline{r_2\,(r_1^2 + r_2^2)}^{k}} + 2\overline{\overline{a_{12}}^{ij}\overline{r_1 r_2}^{k}}
\end{aligned}$$

An alternative to this formulation, which is quite heavy and cumbersome to compute, is that of projecting the modes $\phi^{(x)}$, $\phi^{(y)}$ and $\phi^{(z)}$ and then compute noise, variance and Itô-Stokes drift. Alternatively, one can choose to project only the noise and the Itô-Stokes

drift after its computation, and computing the projected diffusion implicitly applying the double advection procedure to the projected noise.

A.6 Stochastic pressure discretization

The stochastic pressure is defined as

$$\underline{\text{Hydrostatic balance:}}$$

$$p'(x,y,z) = \frac{\nu}{3}\nabla\cdot\boldsymbol{v}_s\Big|_{\eta_b}^{z} + \int_{\eta_b}^{z} b + \boldsymbol{v}_s\cdot\nabla w + \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla w)\,\mathrm{d}\zeta, \tag{A.17}$$

$$\mathrm{d}p_t^{\sigma}(x,y,z) = \int_{\eta_b}^{z} \boldsymbol{\sigma}\mathrm{d}\mathbf{B}_t\cdot\nabla w\,\mathrm{d}\zeta. \tag{A.18}$$

All these terms (except for the divergence of the Itô-Stokes drift and buoyancy) are computed starting from the advection of the vertical velocity w:

$$\mathcal{A}_w = -\left[\overline{\frac{1}{e_{1f}}\overline{u_{\text{adv}}}^{\,k+1/2}\delta_{i+1/2}[w^{\text{in}}]}^{\,k} + \overline{\frac{1}{e_{2t}}\overline{v_{\text{adv}}}^{\,k+1/2}\delta_{j+1/2}[w^{\text{in}}]}^{\,k} + \overline{\frac{1}{e_{3vw}}w_{\text{adv}}\delta_{k+1/2}[w^{\text{in}}]}^{\,k}\right].$$

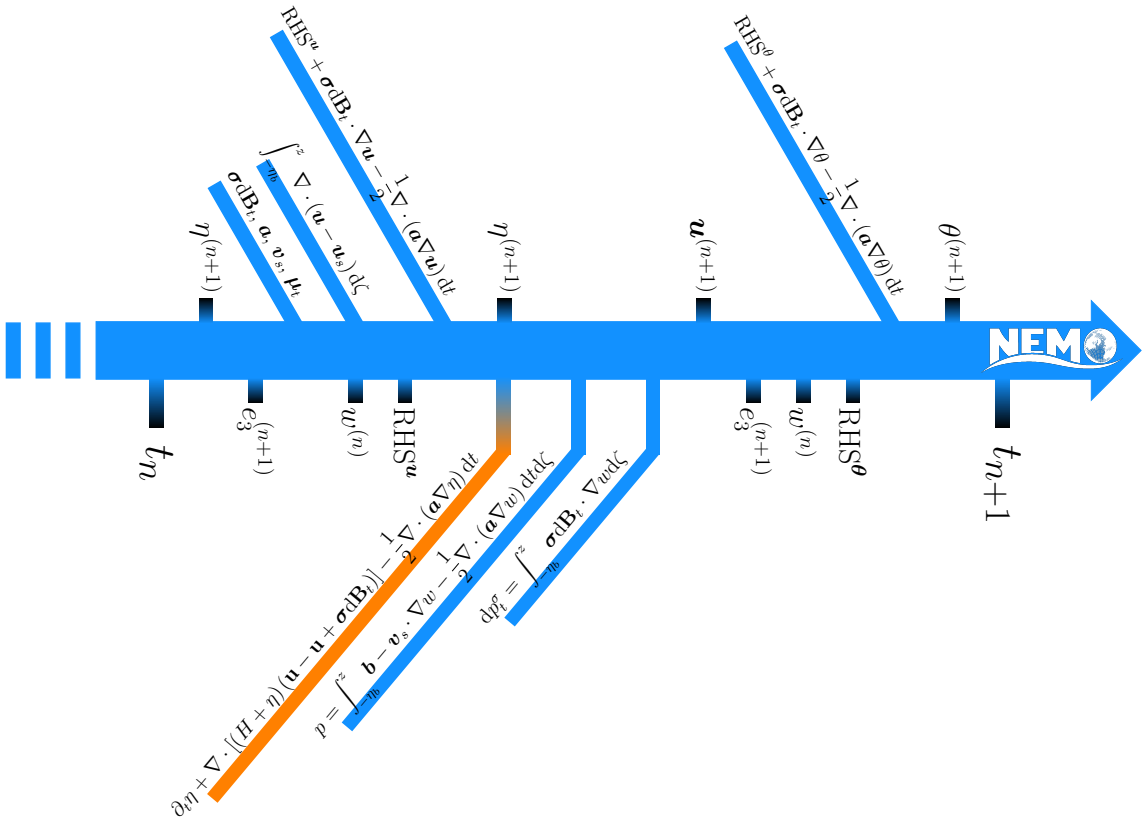
A.7 Coriolis contribution

In the NEMO framework, the Coriolis noise contribution is

$$\mathcal{C}_{\boldsymbol{u}} = -\boldsymbol{f}\times\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t = \begin{bmatrix} +f\,\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{(y)} \\ -f\,\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{(x)} \end{bmatrix} = \begin{bmatrix} +\frac{1}{e_{1u}}\overline{\left(\frac{f}{e_{3f}}\overline{e_{1v}e_{3v}\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{(y)}}^{\,i+1/2}\right)}^{\,j} \\ -\frac{1}{e_{2v}}\overline{\left(\frac{f}{e_{3f}}\overline{e_{1u}e_{3u}\boldsymbol{\sigma}_t\mathrm{d}\mathbf{B}_t^{(x)}}^{\,j+1/2}\right)}^{\,i} \end{bmatrix} \tag{A.19}$$

A.8 Timestepping

The following diagram represents the state of the implementation of Location Uncertainty in NEMO. At the two extremes of the time marching arrow the two time steps t_n and t_{n+1} are present. All the ticks with dark end are computations within NEMO that are left unchanged. The blue appendices are computations that are introduced by the author to implement the method: the diagram represents their algorithmic positioning with respect to the standard operations. The orange appendix describes the barotropic



sub-stepping to compute the sea surface height. It is of a different color as it is not implemented in the current version of the code. However, stochasticity still influences this routine through the baroclinic forcing of the shallow water momentum equations.

BIBLIOGRAPHY

Auclair, F., Rachid Benshila, Lucie Bordois, Martial Boutet, Maurice Brémond, Matthieu Caillaud, Gildas Cambon, Xavier Capet, Laurent Debreu, Nicolas Ducousso, François Dufois, Franck Dumas, Christian Ethé, Jonathan Gula, Christophe Hourdin, Serena Illig, Swen Jullien, Matthieu Le Corre, Solène Le Gac, Sylvie Le Gentil, Florian Lemarié, Patrick Marchesiello, Camille Mazoyer, Guillaume Morvan, Cyril Nguyen, Pierrick Penven, Renaud Person, Joris Pianezze, Stéphane Pous, Lionel Renault, Laurent Roblou, Andres Sepulveda, and Sebastien Theetten (Dec. 2022), *Coastal and Regional Ocean COmmunity model*, version 1.3, DOI: 10.5281/zenodo.7415343.

Bailly, C. and G. Comte-Bellot (2015), *Turbulence*, Springer.

Batchelor, G. K. (2000), *An Introduction to Fluid Dynamics*, Cambridge Mathematical Library, Cambridge University Press.

Bauer, W., P. Chandramouli, L. Li, and E. Mémin (2020a), « Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach », *in*: *Journal of Physical Oceanography*.

— (2020b), « Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models », *in*: *Ocean Modelling* 151.

Bennett, A. F. (2002), *Inverse modeling of the ocean and atmosphere*, Cambridge, UK ; Cambridge University Press, xxii, 234 p.

Berloff, P. S. (2005), « Random-forcing model of the mesoscale oceanic eddies », *in*: *Journal of Fluid Mechanics* 529, pp. 71–95.

Blumberg, A. F. and G. L. Mellor (1987), « A Description of a Three-Dimensional Coastal Ocean Circulation Model », *in*: *Three-Dimensional Coastal Ocean Models*, pp. 1–16.

Borisenko, A. I., I. E. Tarapov, and R. A. Silverman (1979), *Vector and Tensor Analysis with Applications*, Dover Publications.

Bougeault, P. and P. Lacarrère (1989), « Parametrization of Orography-Induced Turbulence in a Mesobeta-Scale Model », *in*: *Monthly Weather Review* 117.*8*, pp. 1872–1890.

Boulvard, P.M. and E. Mémin (2023), « Diagnostic of the Lévy area for geophysical flow models in view of defining high order stochastic discrete-time schemes », *in*: *Foundations of Data Science*.

Boussinesq, J. V. (1903), *Théorie analytique de la chaleur*, Paris: Gauthier-Villars.

Brecht, R., L. Li, W. Bauer, and E. Mémin (2021), *Rotating shallow water flow under location uncertainty with a structure-preserving discretization.*

Brémaud, P. (2020), « Probability Theory and Stochastic Processes », *in*: *Universitext.*

Bricaud, C., J. Le Sommer, G. Madec, C. Calone, J. Deshayes, C. Ethe, J. Chanut, and M. Levy (2020), « Multi-grid algorithm for passive tracer transport in the NEMO ocean circulation model: a case study with the NEMO OGCM (version 3.6) », *in*: *Geoscientific Model Development* 13, pp. 5465–5483.

Bryan, K. (1967), « A numerical method for the study of the circulation of the World Ocean », *in.*

Bryan, K. and M. D. Cox (1968), « A Nonlinear Model of an Ocean Driven by Wind and Differential Heating: Part I. Description of the Three-Dimensional Velocity and Density Fields », *in*: *Journal of the Atmospheric Sciences* 25.*6*, pp. 945–967.

Callies, J., R. Ferrari, J. M. Klymak, and J. Gula (2015), « Seasonality in submesoscale turbulence », *in*: *Nature Communications* 6.

Canuto, C., M. Y. Hussaini, A. Quarteroni, and T. A. Jr. Zang (1988), *Spectral Methods in Fluid Dynamics*, 1st ed., Springer Series in Computational Physics, Springer-Verlag Berlin Heidelberg.

Capet, X., J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin (2008), « Mesoscale to Submesoscale Transition in the California Current System. Part I: Flow Structure, Eddy Flux, and Observational Tests », *in*: *Journal of Physical Oceanography* 38, pp. 29–43.

Carcione, José M. (1999), « Staggered mesh for the anisotropic and viscoelastic wave equation », *in*: *GEOPHYSICS* 64.*6*, pp. 1863–1866.

Cardin, F. and M. Favretti (2013), *Modelli fisico matematici*, Scienze matematiche, fisiche e naturali, Cooperativa Libraria Editrice Università di Padova (CLEUP).

Chandramouli, P., E. Mémin, and D. Heitz (2020), « 4D large scale variational data assimilation of a turbulent flow with a dynamics error model », *in*: *Journal of Computational Physics* 412.

Chandramouli, P., E. Mémin, D. Heitz, and S. Laizet (2018), « Coarse large-eddy simulations in a transitional wake flow with flow models under location uncertainty », *in*: *Computers & Fluids* 168, pp. 170–189.

Chapron, B., P. Dérian, E. Mémin, and V. Resseguier (2018), « Large-scale flows under location uncertainty: a consistent stochastic framework », *in*: *Quarterly Journal of the Royal Meteorological Society* 144, pp. 251–260.

Charney, J. G. (1951), « Dynamic Forecasting by Numerical Process », *in*: *Compendium of Meteorology: Prepared under the Direction of the Committee on the Compendium of Meteorology*, pp. 470–482.

Chatterjee, A. (2000), « An introduction to the proper orthogonal decomposition », *in*: *Current Science* 78.7, pp. 808–817.

Cintolesi, C. and E. Mémin (2020a), « Stochastic Modelling of Turbulent Flows for Numerical Simulations », *in*: *Fluids* 5, p. 108.

— (2020b), « Stochastic modelling of turbulent flows for numerical simulations », unpublished.

Cordier, L and M Bergmann (2003), « Post-Processing of experimental and numerical data », *in*.

Cordier, L. and Michel Bergmann (2008), « Proper Orthogonal Decomposition: an overview », *in*.

Cox, M. D. (1975), « A baroclinic numerical model of the world ocean: Preliminary results », *in*.

Craik, A. D. D. and S. Leibovich (1976), « A rational model for Langmuir circulations », *in*: *Journal of Fluid Mechanics* 73.3, pp. 401–426.

Da Prato, G. and J. Zabczyk (2014), *Stochastic Equations in Infinite Dimensions*, 2nd ed, Encyclopedia of mathematics and its applications / Cambridge university press 152, Cambridge University Press.

Daubechies, Ingrid (1992), *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics.

De Zolt, S., P. Lionello, A. Nuhu, and A. Tomasin (2006), « The disastrous storm of 4 November 1966 on Italy », *in*: *Natural Hazards and Earth System Science* 6, pp. 861–879.

Debreu, L. and E. Blayo (2002), *AGRIF: Adaptive Grid Refinement In Fortran*, Research Report RT-0262, INRIA, p. 16, URL: `https://inria.hal.science/inria-00069912`.

— (2008), « Two-way embedding algorithms: a review », *in*: *Deutsche Hydrographische Zeitschrift* 58, pp. 415–428.

Debreu, L., P. Marchesiello, P. Penven, and G. Cambon (2012), « Two-way nesting in split-explicit ocean models: Algorithms, implementation and validation », *in*: *Ocean Modelling* 49-50, pp. 0–21.

Debussche, A., B. Hug, and E. Mémin (2023), « A Consistent Stochastic Large-Scale Representation of the Navier–Stokes Equations », *in*: *Journal of Mathematical Fluid Mechanics* 25.*1*.

Defina, L. D'Alpaos; A. (2007), « Mathematical modeling of tidal hydrodynamics in shallow lagoons: A review of open issues and applications to the Venice lagoon », *in*: *Computers & Geosciences* 33, pp. 476–496.

Della Penna, A. and P. Gaube (2019), « Overview of (Sub)mesoscale Ocean Dynamics for the NAAMES Field Program », *in*: *Frontiers in Marine Science* 6, p. 384.

Demange, J., L. Debreu, P. Marchesiello, F. Lemarié, E. Blayo, and C. Eldred (2019), « Stability analysis of split-explicit free surface ocean models: Implication of the depth-independent barotropic mode approximation », *in*: *Journal of Computational Physics* 398.

Dukowicz, J. K. and R. D. Smith (1994), « Implicit free-surface method for the Bryan-Cox-Semtner ocean model », *in*: *Journal of Geophysical Research: Oceans* 99, pp. 7991–8014.

Dumas, F., R. Le Gendre, Y. Thomas, and S. Andréfouët (2012), « Tidal flushing and wind driven circulation of Ahe atoll lagoon (Tuamotu Archipelago, French Polynesia) from in situ observations and numerical modelling », *in*: *Marine Pollution Bulletin* 65 (10–12).

Durand, M., L.-L. Fu, D. P. Lettenmaier, D. E. Alsdorf, E. Rodriguez, and D. Esteban-Fernandez (2010), « The Surface Water and Ocean Topography Mission: Observing Terrestrial Surface Water and Oceanic Submesoscale Eddies », *in*: *Proceedings of the IEEE* 98.*5*, pp. 766–779, (visited on 12/11/2019).

Durran, D., J. A. Weyn, and M. Q. Menchaca (2017), « Practical Considerations for Computing Dimensional Spectra from Gridded Data », *in*: *Monthly Weather Review* 145, pp. 3901–3910.

Dynamic Mode Decomposition (2016), Philadelphia, PA: Society for Industrial and Applied Mathematics.

Fasshauer, G. E. (2007), *Meshfree Approximation Methods with Matlab*, World SCientific.

Fiorini, C., P.M. Boulvard, L. Li, and Mémin (2022), « A Two-Step Numerical Scheme in Time for Surface Quasi Geostrophic Equations Under Location Uncertainty », *in*: pp. 57–67.

Flandoli, F. and E. Luongo (2023), *Stochastic Partial Differential Equations in Fluid Mechanics*, Lecture Notes in Mathematics, 2328, Springer.

Fox-Kemper, B., A. Adcroft, C. W. Böning, E. P. Chassignet, E. Curchitser, G. Danabasoglu, C. Eden, M. H. England, R. Gerdes, R. J. Greatbatch, S. M. Griffies, R. W. Hallberg, E. Hanert, P. Heimbach, H. T. Hewitt, C. N. Hill, Y. Komuro, S. Legg, J. Le Sommer, S. Masina, S. J. Marsland, S. G. Penny, F. Qiao, T. D. Ringler, A. M. Treguier, H. Tsujino, P. Uotila, and S. G. Yeager (2019), « Challenges and Prospects in Ocean Circulation Models », *in*: *Frontiers in Marine Science* 6, p. 65, (visited on 11/29/2019).

Fox-Kemper, Baylor, Alistair Adcroft, Claus W. Böning, Eric P. Chassignet, Enrique Curchitser, Gokhan Danabasoglu, Carsten Eden, Matthew H. England, Rüdiger Gerdes, Richard J. Greatbatch, Stephen M. Griffies, Robert W. Hallberg, Emmanuel Hanert, Patrick Heimbach, Helene T. Hewitt, Christopher N. Hill, Yoshiki Komuro, Sonya Legg, Julien Le Sommer, Simona Masina, Simon J. Marsland, Stephen G. Penny, Fangli Qiao, Todd D. Ringler, Anne Marie Treguier, Hiroyuki Tsujino, Petteri Uotila, and Stephen G. Yeager (2019), « Challenges and Prospects in Ocean Circulation Models », *in*: *Frontiers in Marine Science* 6, p. 65.

Franzke, C., A. J. Majda, and E. Vanden-Eijnden (2005), « Low-Order Stochastic Mode Reduction for a Realistic Barotropic Model Climate », *in*: *Journal of the Atmospheric Sciences* 62.*6*, pp. 1722–1745, (visited on 11/29/2019).

Gent, P. R. and J. C. Mcwilliams (1990), « Isopycnal Mixing in Ocean Circulation Models », *in*: *Journal of Physical Oceanography* 20, pp. 150–155.

Gent, Peter R. (2011), « The Gent–McWilliams parameterization: 20/20 hindsight », *in*: *Ocean Modelling* 39, pp. 2–9.

Grabowski, W. W. and P. K. Smolarkiewicz (1999), « CRCP: a Cloud Resolving Convection Parameterization for modeling the tropical convecting atmosphere », *in*: *Physica D*, p. 8.

Gray, D. D. and A. Giorgini (1976), « The validity of the boussinesq approximation for liquids and gases », *in*: *International Journal of Heat and Mass Transfer* 19, pp. 545–551.

Griffies, S. and A. Adcroft (2008), « Formulating the Equations of Ocean Models », *in*: *Eddy resolving ocean models, Geophysical Monograph* 177, pp. 281–317.

Griffies, S. M., C. Böning, F. O. Bryan, E. P. Chassignet, R. Gerdes, H. Hasumi, A. Hirst, A.-M. Treguier, and D. Webb (2000), « Developments in ocean climate modelling », *in*: *Ocean Modelling* 2, pp. 123–192.

Griffies, S. M. and A.-M. Treguier (2013), « Ocean Circulation Models and Modeling », *in*: *Ocean Circulation and Climate*, vol. 103, International Geophysics, Academic Press, chap. 20, pp. 521–551.

Grooms, I., A. Majda, and S. Smith (2014), « Stochastic superparameterization in a quasigeostrophic model of the Antarctic Circumpolar Current », *in*: *Ocean Modelling* 85.

Grooms, I. and A. J. Majda (2013), « Efficient stochastic superparameterization for geophysical turbulence », *in*: *Proceedings of the National Academy of Sciences* 110.*12*, pp. 4464–4469.

Harouna, S. K. and E. Mémin (2017), « Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling », *in*: *Computers & Fluids* 156, pp. 456–469.

Heinrich, J. C., P. S. Huyakorn, O. C. Zienkiewicz, and A. R. Mitchell (1977), « An 'upwind' finite element scheme for two-dimensional convective transport equation », *in*: *International Journal for Numerical Methods in Engineering* 11, pp. 131–143.

Holm, D. D. (2015), « Variational principles for stochastic fluid dynamics », *in*: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 471.

Holmes, P., Lumley J. L., and G. Berkooz (1996), *Turbulence, coherence structures, dynamical systems and symetry*, Cambridge university press.

Holmes, P., J. L. Lumley, and G. Berkooz (1996), *Turbulence, coherent structures, dynamical systems, and symmetry*, Cambridge monographs on mechanics, Cambridge University Press.

Huang, R. X. (1993), « Real Freshwater Flux as a Natural Boundary Condition for the Salinity Balance and Thermohaline Circulation Forced by Evaporation and Precipitation », *in*: *Journal of Physical Oceanography* 23, pp. 2428–2446.

Hurlburt, H. E. and P. J. Hogan (2000), « Impact of 1/8° to 1/64° resolution on Gulf Stream model–data comparisons in basin-scale subtropical Atlantic Ocean models », en, *in*: *Dynamics of Atmospheres and Oceans* 32, pp. 283–329.

Idier, D., F. Dumas, and H. Muller (2012), « Tide-surge interaction in the English Channel », *in*: *Natural Hazards and Earth System Science* 12, pp. 3709–3718.

Ilıcak, M., A. J. Adcroft, S. M. Griffies, and R. W. Hallberg (2012), « Spurious dianeutral mixing and the role of momentum closure », *in*: *Ocean Modelling* 45–46, pp. 0–58.

Jamet, Q., E. Mémin, F. Dumas, L. Li, and P. Garreau (2023), « Toward a Stochastic Parameterization for Oceanic Deep Convection », *in*: *Stochastic Transport in Upper Ocean Dynamics II*, Springer International Publishing, pp. 143–157.

Jouini, M., M. Levy, M. Crépon, and S. Thiria (2013), « Reconstruction of satellite chlorophyll images under heavy cloud coverage using a neural classification method », *in*: *Remote Sensing of Environment*, pp. 232–246.

Kalnay, E. (2012), *Atmospheric modeling, data assimilation and predictability*, 7. print, Cambridge: Cambridge University Press.

Kerschen, G., J.-C. Golinval, A. F. Vakakis, and L. A. Bergman (2005), « The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview », *in*: *Nonlinear Dynamics* 41, pp. 147–169.

Killworth, P. D., D. J. Webb, D. Stainforth, and S. M. Paterson (1991), « The Development of a Free-Surface Bryan–Cox–Semtner Ocean Model », *in*: *Journal of Physical Oceanography* 21, pp. 1333–1348.

Kolmogorov, A. N. (1941), « The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers », *in*: *Proceedings: Mathematical and Physical Sciences* 434, pp. 9–13.

Kolmogorov, A. N., S. V. Fomin, and R. A. Silverman (1975), *Introductory Real Analysis*, 1st, Dover Publications.

Kundu, P. K., I. M. Cohen, and D. R. Dowling (2015), *Fluid Mechanics*, 6th ed., Academic Press.

Kunisch, K. and S. Volkwein (2008), « Proper orthogonal decomposition for optimality systems », *in*: *ESAIM: Mathematical Modelling and Numerical Analysis* 42.*1*, pp. 1–23.

Kunita, H. (1997), *Stochastic Flows and Stochastic Differential Equations*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

Kutz, J. Nathan, Xing Fu, and Steven L. Brunton (2016), « Multiresolution Dynamic Mode Decomposition », *in*: *SIAM Journal on Applied Dynamical Systems* 15, pp. 713–735.

Landau, L. D. and J. B. Sykes (1987), *Fluid mechanics Volume 6*, 2nd ed., 2nd English ed., rev, Course of theoretical physics 6, Pergamon Press.

Lazure, P. and F. Dumas (2008), « An external–internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS) », *in*: *Advances in Water Resources* 31, pp. 233–250.

Lemarié, F., L. Debreu, G. Madec, J. Demange, J.M. Molines, and M. Honnorat (2015), « Stability constraints for oceanic numerical models: implications for the formulation of time and space discretizations », *in*: *Ocean Modelling* 92, pp. 124–148.

Lévy, M., D. Iovino, L. Resplandy, P. Klein, G. Madec, A.-M. Tréguier, S. Masson, and K. Takahashi (2012), « Large-scale impacts of submesoscale dynamics on phytoplankton: Local and remote effects », *in*: *Ocean Modelling* 43-44, pp. 77–93.

Lévy, M., P. Klein, A.-M. Tréguier, D. Iovino, G. Madec, S. Masson, and K. Takahashi (2010), « Modifications of gyre circulation by sub-mesoscale physics », *in*: *Ocean Modelling* 34, pp. 1–15.

Lévy, M., L. Resplandy, P. Klein, X. Capet, D. Iovino, and C. Ethé (2012), « Grid degradation of submesoscale resolving ocean models: Benefits for offline passive tracer transport », *in*: *Ocean Modelling* 48, pp. 1–9.

Lewy, H., K. Friedrichs, and R. Courant (1928), « Über die partiellen Differenzengleichungen der mathematischen Physik », *in*: *Mathematische Annalen* 100, pp. 32–74.

Li, L. (2021), « Stochastic modeling and numerical simulation of ocean dynamics », PhD thesis, Université de Rennes 1.

Li, L., B. Deremble, N. Lahaye, and E. Mémin (2023), « Stochastic Data-Driven Parameterization of Unresolved Eddy Effects in a Baroclinic Quasi-Geostrophic Model », *in*: *Journal of Advances in Modeling Earth Systems* 15.

Li, L., E. Mémin, and G. Tissot (2022), « Stochastic Parameterization with Dynamic Mode Decomposition », *in*: pp. 179–193.

Loeve, M. (1978), *Probability Theory II*, F.W.Gehring P.r.Halmos and C.c.Moore, Springer.

Lumley, J. L. (1970), *Stochastic Tools in Turbulence*, Applied Mathematics and Mechanics 12, Academic Press, Elsevier.

Madec, G., R. Bourdallé-Badie, J. Chanut, E. Clementi, A. Coward, C. Ethé, D. Iovino, D. Lea, C. Lévy, T. Lovato, N. Martin, S. Masson, S. Mocavero, C. Rousset, D. Storkey, M. Vancoppenolle, S. Müeller, G. Nurser, M. Bell, and G. Samson (Oct. 2019), *NEMO ocean engine*, version v4.0, Add SI3 and TOP reference manuals, DOI: 10.5281/zenodo.3878122.

Mallat, S. (1999), *A Wavelet Tour of Signal Processing*, 2nd ed., Wavelet Analysis & Its Applications, Academic Press.

Maltrud, M. E and J. L McClean (2005), « An eddy resolving global 1/10° ocean simulation », *in*: *Ocean Modelling* 8.*1-2*, pp. 31–54.

Manabe, S. and K. Bryan (1969), « Climate Calculations with a Combined Ocean-Atmosphere Model », *in*: *Journal of the Atmospheric Sciences* 26, pp. 786–789.

Marchesiello, P., X. Capet, C. Menkes, and S. C. Kennan (2011), « Submesoscale dynamics in tropical instability waves », *in*: *Ocean Modelling* 39, pp. 0–46.

McDougall, T. and P.M. Barker (2011), « Getting started with TEOS-10 and the Gibbs Seawater. GSW Oceanographic Toolbox », *in*: *SCOR/IAPSO WG*, pp. 1–28.

McWilliams, J. C. (1996), « Modeling the Oceanic General Circulation », *in*: *Annual Review of Fluid Mechanics* 28, pp. 215–248.

— (2016), « Submesoscale currents in the ocean », *in*: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 472, pp. 1–32.

Mel, R., A. Sterl, and P. Lionello (2013), « High resolution climate projection of storm surge at the Venetian coast », *in*: *Natural Hazards and Earth System Sciences* 13, pp. 1135–1142.

Mémin, E. (2014), « Fluid flow dynamics under location uncertainty », *in*: *Geophysical & Astrophysical Fluid Dynamics* 108, pp. 119–146.

Mikulevicius, R. and B. L. Rozovskii (2004), « Stochastic Navier–Stokes Equations for Turbulent Flows », *in*: *SIAM Journal on Mathematical Analysis* 35, pp. 1250–1310.

Murray, R. J. and C.J.C. Reason (2001), « A Curvilinear Version of the Bryan–Cox–Semtner Ocean Model and Its Representation of the Arctic Circulation », *in*: *Journal of Computational Physics* 171, pp. 1–46.

Oertel, H. (2010), *Prandtl-Essentials of Fluid Mechanics*, 3rd ed., Applied Mathematical Sciences (Switzerland), Springer.

Olivella, X. O. and C. Agelet de Saracibar Bosch (2000), *Mecanica de Medios Continuos para Ingenieros*, Politext.

Paiva, A. M., E. P. Chassignet, and A. J. Mariano (2000), « Numerical simulations of the North Atlantic subtropical gyre: sensitivity to boundary conditions », *in*: *Dynamics of Atmospheres and Oceans* 32, pp. 209–237.

Pedlosky, J. (1990), *Geophysical Fluid Dynamics*, 2nd, Springer.

— (1996), *Ocean Circulation Theory*, 1st ed., Springer.

Pinier, B., E. Mémin, S. Laizet, and R. Lewandowski (2019), « Stochastic flow approach to model the mean velocity profile of wall-bounded flows », *in*: *Physical Review E* 99.

Pope, S. B. (2000), *Turbulent flows*, 1st ed., Cambridge University Press.

Porta Mana, P. and L. Zanna (2014), « Toward a stochastic parameterization of ocean mesoscale eddies », *in*: *Ocean Modelling* 79, pp. 1–20.

Prandtl, L. (1925), « Über die ausgebildeten Turbulenz », *in*: *Z. Angew. Math. Mech.* 5, pp. 136–139.

Protter, P. E. (2004), *Stochastic integration and differential equations*, 2nd ed, Applications of mathematics, Springer.

Quarteroni, A. and A. Valli (1994), *Numerical approximation of partial differential equations*, 1st ed., Springer Series in Computational Mathematics №23, Springer.

Redi, M. H. (1982), « Oceanic Isopycnal Mixing by Coordinate Rotation », *in*: *Journal of Physical Oceanography* 12, pp. 1154–1158.

Reeks, M.W. (1983), « The transport of discrete particles in inhomogeneous turbulence », *in*: *Journal of Aerosol Science* 14, pp. 0–739.

Resseguier, V., L. Li, G. Jouan, P. Derian, E. Mémin, and B. Chapron (2021), « New Trends in Ensemble Forecast Strategy: Uncertainty Quantification for Coarse-Grid Computational Fluid Dynamics », *in*: *Archives Of Computational Methods In Engineering* 28, ed. by Springer Science, pp. 215–261.

Resseguier, V., E. Mémin, and B. Chapron (2017a), « Geophysical flows under location uncertainty, Part I Random transport and general models », *in*: *Geophysical & Astrophysical Fluid Dynamics* 111, pp. 149–176.

— (2017b), « Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading », *in*: *Geophysical & Astrophysical Fluid Dynamics* 111, pp. 177–208.

— (2017c), « Geophysical flows under location uncertainty, Part III SQG and frontal dynamics under strong turbulence conditions », *in*: *Geophysical & Astrophysical Fluid Dynamics* 111, pp. 209–227.

Resseguier, V., E. Mémin, D. Heitz, and B. Chapron (2017), « Stochastic modelling and diffusion modes for proper orthogonal decomposition models and small-scale flow analysis », *in*: *Journal of Fluid Mechanics* 826, pp. 888–917.

Reynolds, O. (1883), « An Experimental Investigation of the Circumstances Which Determine Whether the Motion of Water Shall Be Direct or Sinuous, and of the Law of

Resistance in Parallel Channels », *in*: *Philosophical Transactions of the Royal Society of London* 174, pp. 935–982.

Rivlin, T. J. (1969), *An introduction to the approximation of functions*, 1st ed., Blaisdell Book in Numerical Analysis and Computer Science, Blaisdell Pub. Co.

Roquet, F., G. Madec, T. J. McDougall, and P. M. Barker (2015), « Accurate polynomial expressions for the density and specific volume of seawater using the TEOS-10 standard », *in*: *Ocean Modelling* 90, pp. 29–43.

Rowley, C. (2005), « Model reduction for fluids, using balanced proper orthogonal decomposition », *in*: *International Journal of Bifurcation and Chaos* 15.

Rowley, C. W. (2002), « Modeling, Simulation, and Control of Cavity Flow Oscillations », PhD thesis, California Institute of Technology.

Rowley, C. W., I. Mezi, S. Bagheri, P. Schlatter, and D. S. Henningson (2009), « Spectral analysis of nonlinear flows », *in*: *Journal of Fluid Mechanics* 641, pp. 115–127.

Sacco, F., R. Ostilla-Mónico, and R. Verzicco (2020), « Dynamic mode decomposition analysis of coherent structures in rotating plane Couette flow », *in*: *Journal of Physics: Conference Series* 1522, p. 012012.

Sanderson, B. G. (1998), « Order and Resolution for Computational Ocean Dynamics », *in*: *Journal of Physical Oceanography* 28, pp. 1271–1286.

Sasaki, Y. N., S. Minobe, N. Schneider, T. Kagimoto, M. Nonaka, and H. Sasaki (2008), « Decadal Sea Level Variability in the South Pacific in a Global Eddy-Resolving Ocean Model Hindcast », *in*: *Journal of Physical Oceanography* 38.*8*, pp. 1731–1747.

Schmid, P. and J. Sesterhenn (2008), « Dynamic Mode Decomposition of numerical and experimental data », *in*: *Journal of Fluid Mechanics* 656, pp. 5–28.

Semtner, A. J. and R. M. Chervin (1988), « A simulation of the global ocean circulation with resolved eddies », *in*: *Journal of Geophysical Research: Oceans* 93, pp. 15502–15522.

Shchepetkin, A. F. and J. C. McWilliams (2005), « The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model », *in*: *Ocean Modelling* 9, pp. 0–404.

— (2009), « Computational Kernel Algorithms for Fine-Scale, Multiprocess, Longtime Oceanic Simulations », *in*: vol. 14, Elsevier, pp. 121–183.

Siegel, A., J. B. Weiss, J. Toomre, J. C. McWilliams, Pavel S. Berloff, and I. Yavneh (2001), « Eddies and vortices in ocean basin dynamics », *in*: *Geophysical Research Letters* 28.*16*, pp. 3183–3186.

Smagorjnsky, J. (1983), « The beginnings of numerical weather prediction and general circulation modeling: early recollections », *in*: vol. 25, Elsevier, pp. 3–37.

Spiegel, E. A. and G. Veronis (1960), « On the Boussinesq Approximation for a Compressible Fluid. », *in*: *The Astrophysical Journal* 131, p. 442.

Tailleux, R. (2010), « A simple method to construct energetically and thermodynamically consistent Boussinesq approximations with arbitrary nonlinear equations of state », *in*.

Teramoto, T. (1992), *Deep Ocean Circulation*.

Thuburn, J., J. Kent, and N. Wood (2014), « Cascades, backscatter and conservation in numerical models of two-dimensional turbulence », *in*: *Quarterly Journal of the Royal Meteorological Society* 140, pp. 626–638.

Tissot, G., A. Cavalieri, and E. Mémin (2021), « Stochastic linear modes in a turbulent channel flow », *in*: *J. Fluid Mech.* 912.

Tissot, G., E. Mémin, and Q. Jamet (2023), « Stochastic Compressible Navier–Stokes Equations Under Location Uncertainty », *in*: pp. 239–319.

Tu, J., C. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz (2014), « On dynamic mode decomposition: Theory and applications », *in*: *Journal of Computational Dynamics* 1, pp. 391–421.

Tucciarone, F. L., E. Mémin, and L. Li (2022), « Primitive Equations Under Location Uncertainty: Analytical Description and Model Development », *in*: *Stochastic Transport in Upper Ocean Dynamics*, Springer International Publishing, pp. 287–300.

— (2023), « Data driven stochastic primitive equations with dynamic modes decomposition », *in*: *Stochastic Transport in Upper Ocean Dynamics II*, Springer International Publishing, pp. 321–336.

Vallis, G. K. (2017), *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*, 2nd, Cambridge University Press.

Wu, Z., S. L. Brunton, and S. Revzen (2021), « Challenges in dynamic mode decomposition », *in*: *Journal of The Royal Society Interface* 18.*185*, p. 20210686.

Wunsch, C. (2015), *Modern Observational Physical Oceanography. Understanding the Global Ocean*.

Zeytounian, R. K. (2003), « Joseph Boussinesq and his approximation: a contemporary view », *in*: *Comptes Rendus Mécanique* 331, pp. 575–586.

Zienkiewicz, O. C., R. L. Taylor, and P. Nithiarasu (2013), *The Finite Element Method for Fluid Dynamics*, 7th ed., Butterworth-Heinemann.

COLLEGE DOCTORAL BRETAGNE
MATHS, TELECOMS
INFORMATIQUE, SIGNAL
SYSTEMES, ELECTRONIQUE

Université de Rennes

Titre : Paramétrisation stochastique des modéles océaniques à l'aide d'observations à haute résolution

Mot clés : Paramétrisation stochastique, Modélisation des océans, Équations primitives, Mécanique des fluides, NEMO

Résumé : Le climat dépend fortement de l'état global de l'océan. La simulation numérique reste le seul moyen de prévoir le système océan-atmosphère et d'évaluer ses états futurs afin d'établir des prévisions fiables des évolutions météorologiques et climatologiques. Les simulations à grande échelle constituent le principaux outils d'étude de l'océan et de l'atmosphère, dans mesure où les simulations à haute résolution restent confinées à de petits domaines géographiques ou à de courtes périodes d'intégration. L'interdépendance complexe des dynamiques à méso-échelle et sous-méso-échelle est cependant perdue dans les simulations qui ne résolvent pas les échelles inférieures au rayon de déformation de Rossby ; celles-ci doivent donc être paramétrées. La plupart des défis associés à la dynamique des fluides (dans toutes ses connotations) découlent de la représentation de ces effets à l'aide d'un schéma de fermeture efficace. Une nouvelle famille d'approches consiste à incorporer des perturbations et des composantes de bruit dans la dynamique. L'objectif est d'enrichir la variabilité et de paramétrer les processus sous-maille, la turbulence, l'incertitude des valeurs limites et de tenir compte des erreurs numériques et de discrétisation, tout en respectant les principes physiques de la dynamique des fluides.

Title: Stochastic parametrization of ocean models through high resolution observations

Keywords: Stochastic parametrization, Ocean modelling, Primitive Equations, Fluid Mechanics, NEMO

Abstract: The global climate is strongly dependent on the global Ocean state. Numerical simulation remains the only way to forecast the Ocean-Atmosphere system and assess future states to make reliable meteorological and climatological hazard forecasts. The primary tool for the investigation of the Ocean and the Atmosphere are large-scale simulations, while high resolution simulations remains confined to small geographical domains or short integration periods. The complex interdependence of mesoscale and submesoscale dynamics is, however, lost in simulations that do not resolve scales below the Rossby radius of deformation, and thus must be parametrized. Most of the challenges of fluid dynamics (in all its connotations) arise from the representation of these effects with an efficient closure scheme. A novel research trend involves incorporating perturbations and noise components into the dynamics. The goal is to enrich the variability and parametrize subgrid processes, turbulence, boundary value uncertainty and account for numerical and discretization errors.