Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach

Werner Bauer, Pranav Chandramouli, Bertrand Chapron, Long Li, Etienne Mémin

HAL Id: hal-02398521
https://hal.inria.fr/hal-02398521

Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach

Werner Bauer[1], Pranav Chandramouli[1], Bertrand Chapron[2], Long Li [1] and Etienne Mémin [1*]

[1]Inria/ IRMAR, Campus universitaire de Beaulieu, 35042 Rennes Cedex, France

[2]Ifremer, LOPS, Pointe du Diable, Plouzané 29280, France

Abstract

An important open question in fluid dynamics concerns the effect of small-scales in structuring a fluid flow. In oceanic or atmospheric flows, this is aptly captured in wave-current interactions through the study of the well-known Langmuir secondary circulation. Such wave-current interactions are described by the Craik-Leibovich system, in which the action of a wave induced velocity, the Stokes drift, produces a so called "vortex force" that causes streaking in the flow. In this work, we show that these results can be generalized as a generic effect of the spatial inhomogeneity of the statistical properties of the small-scale flow components. As demonstrated, this is well captured through a stochastic representation of the flow.

1 Introduction

The in-depth understanding of the interactions between large-scale oceanic currents and small-scale fluctuations of the velocity is of utmost importance for numerical simulation and analysis of oceanic flows. Today, and for the foreseeable future, small-scale velocity fluctuations and their contributions to the large-scale flow cannot be fully resolved in numerical simulations. These small-scale fluctuations arising from phenomena such as surface gravity waves, internal waves, or turbulence, interact with the large-scale flow in a nonlinear way. These interactions and their effects must be adequately modeled. A purely diffusive model, describing the mixing effects of the small-scale components on the resolved system's variables, will generally fail to represent the emergence of secondary circulations such as Langmuir cells or velocity streaks in wall-bounded turbulent flows.

The interaction of surface gravity waves with the mean current has been a subject of intense investigations since the early seventies (Ardhuin et al. 2008; Craik and Leibovich 1976; Hasselmann 1971; Holm 1996; Leibovich 1980; Longuet-Higgins and Stewart 1964; Phillips 1977). These studies have focused on the nonlinear interactions between wave orbital velocities and the upper mean flow. In particular, they focus on the role of the residual Stokes drift as the main driver of this interaction. Two parallel representations are presented in literature: in the first the driving force is attributed to the action of the so-called "vortex force" while in the second, it is attributed to a radiation stress (Longuet-Higgins and Stewart 1964; Mellor 2003). These two representations are related (Lane et al. 2007) although both have been used to understand different physical phenomena: wave setup in the coastal surf zone, long-shore currents, infra-gravity waves, for the latter and Langmuir circulation for the former. The vortex force and its associated Bernoulli head pressure term is now commonly used in oceanic simulation to take into account the effect of surface gravity waves (Harcourt and D'Asaro 2008; Harcourt 2015; McWilliams et al. 1997; Teixeira and Belcher 2010). Several equivalent forms of these wave-averaged representations are collated and analyzed in (Suzuki and Fox-Kemper 2016).

It is important to note that the validity of these models are predicated on several assumptions. There is a separation assumption on the temporal scales between the waves and the circulation. In addition, the steepness of the waves is assumed to be small. The waves are also assumed to be quasi-irrotational with a divergence free Stokes drift (Leibovich 1980; McWilliams et al. 2004; Mellor 2016). All these models rely on an explicit representation of the waves, a natural requirement given the interactions they aim to represent. However, there is an inherent limitation associated to such a formulation: they can not be easily extended to include other interactions between small-scale velocity fluctuations and mean currents.

The objective of this work is to provide a consistent mathematical framework capable of accounting for more general interactions between unresolved scales and the large-scale flow. This is achieved by a stochastic representation of the flow dynamics (Chapron et al. 2018; Mémin 2014; Resseguier et al. 2017a,b,c,d). In this representation, the velocity field, decomposed into a large scale flow component and a small-scale random component, may reflect an effective advection characterized by the inhomogeneity of the random flow component. This advection correction will be shown to encode within it, the Craik Leibovich vortex

---

**Corresponding author. Email: etienne.memin@inria.fr

force. Thus, this representation, based on less restrictive assumptions, would encompass the original Craik-Leibovich system as a particular instance. Yet, more complex interactions can be captured.

Stochastic representations of geophysical flows have long been the focus of intensive research efforts (Franzke and Majda 2005; Hasselmann 1976). These investigations are motivated by the need to design dynamic systems capable of taking into account error models for data assimilation or ensemble forecasting. For these purposes, several different schemes have been proposed in literature. For example, many early turbulence models were developed principally to quantify the effects of energy backscatter from the fast/small scales of motion to the larger scales (Leith 1990). Related models have focused greater attention on oceanography and weather forecasting (Buizza et al. 1999; Porta Mana and Zanna 2014; Shutts 2005). They have been shown to improve the system forecasting skill for weather prediction but not for oceanic flows (Andrejczuk et al. 2016). In order to cope with the energy dissipation due to unresolved scales, but also to limit the increase in variance, these models are often artificially damped by eddy-viscosity diffusion. This concept, introduced via an analogy with the molecular viscosity mechanism, dates back to the work of Boussinesq (Boussinesq 1877) but has no basis on uncertainty modeling. Thus, a careful tuning of the noise parameter combined with an unrelated viscosity to stabilize the model is required for implementing these models successfully. The success of such tuning methods often do not extend into new flow regimes.

Contrarily, the stochastic framework proposed in this work is derived from physical conservation laws expressed through the stochastic transport of fluid parcels. This ensures energy conservation and provides new approaches to sub-grid parameterization, expressed both in terms of fluctuation distributions, and spatial/temporal correlations. A very related circulation preserving approach derived from Hamiltonian principles has been proposed in (Holm 2015). Energy conserving randomization have also been recently proposed for climate modeling (Gugole and Franzke 2019). As such, these models introduce stochastic corrections that are amenable to statistical inference from data (either observed, or numerical high-resolution data) and yield new data analysis tools and models for turbulent flows (Resseguier et al. 2017d). This class of models establishes a physically relevant randomization via its rigorous derivation at the fundamental level, rather than a perturbation of the initial condition or ad hoc forcing (Chapron et al. 2018).

The paper is organized as follows. In the proceeding section, the Craik-Leibovich theory is briefly recalled (Andrews and McIntyre 1978; Craik and Leibovich 1976; Leibovich 1980; Holm 1996) delineating the role played by the Stokes drift. We then present in Section 3, the stochastic framework applied in this work. In Section 4 we draw a framework parallel between the Craik-Leibovich system and our stochastic framework leading to a general stochastic Craik-Leibovich system. In Section 5, some numerical results are provided for a simple barotropic quasi-geostrophic stochastic system showing the structuring effect of the noise on the large-scale flow. The stochastic model is analyzed, and we demonstrate the global energy conservation property of the model.

2   Craik-Leibovich representation of waves-mean current interactions

The Craik-Leibovich (CL) equations (see (Craik and Leibovich 1976; Holm 1996; Leibovich 1980; McWilliams et al. 1997; Suzuki and Fox-Kemper 2016) for different elegant derivations) parametrize the effect of surface gravity waves on the (large-scale) evolution of the mean current ($\overline{\boldsymbol{u}}$) :

$$\partial_t \overline{\boldsymbol{u}} + (\overline{\boldsymbol{u}} \cdot \boldsymbol{\nabla})\overline{\boldsymbol{u}} = -\boldsymbol{\nabla}\pi + \boldsymbol{u}_S \times \boldsymbol{\omega}, \quad \boldsymbol{\nabla} \cdot \overline{\boldsymbol{u}} = 0, \quad \pi = p + \frac{1}{2}\|\overline{\boldsymbol{u}} + \boldsymbol{u}_S\|^2 - \frac{1}{2}\|\overline{\boldsymbol{u}}\|^2, \tag{1}$$

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

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

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

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

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

$$\overline{\boldsymbol{u}} = \boldsymbol{u}_L - \boldsymbol{u}_S + O(\epsilon^4). \tag{3}$$

The vortex force $\boldsymbol{u}_S \times \boldsymbol{\omega}$ acts as a Lamb vector between the mean flow vorticity and the Stokes drift. The Stokes drift and the associated vortex force term in the momentum equation have been shown to be the main players in the interaction between the gravity waves and the mean current. Note that in its various

derivations, the Craik-Leibovich system requires to assume irrotational waves with a weak slope and a divergence free Stokes drift (Holm 1996; Leibovich 1980; McWilliams et al. 2004). Formally, it does not strictly apply to general interactions between small-scale velocity fluctuations and the resolved flow. In the following, we show how a stochastic representation enables us to model such interactions by expressing them as a generic effect of the spatial inhomogeneity of the statistical properties of the velocity fluctuations leading to the emergence of an associated vortex force.

3   Stochastic flow models under location uncertainty

The stochastic framework developed in this section, termed modeling under location uncertainty (LU), relies on a decomposition in a Lagrangian setting of the velocity in terms of a smooth in time component $\boldsymbol{u}$ and an incompressible highly oscillating random component

$$\frac{\mathrm{d}\boldsymbol{X}_t}{\mathrm{d}t} = \boldsymbol{u}(\boldsymbol{X}_t, t) + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\dot{\boldsymbol{B}}_t. \tag{4}$$

It is important to outline, that contrary to traditional Large Eddy Simulation (LES) settings, this decomposition corresponds to a temporal decomposition and not a spatial decomposition formulated through spatial filters and/or decimation operators. The resolved velocity component $\boldsymbol{u}$ corresponds to a smooth Lagrangian quantity. The second term, referred to as the random component, is written (formally) as the time derivative of a d-dimensional Brownian function $\dot{\boldsymbol{B}}_t = \mathrm{d}\boldsymbol{B}_t/\mathrm{d}t$. It represents the fast, unresolved velocity component. In turbulent flows, time and spatial scales are related, and in the inertial range the turn-over time ratio for two different scales, $\tau_L/\tau_\ell \propto (L/\ell)^{2/3}$, exhibits a direct relation between a change of time scale and a change of spatial resolution. A coarsening in time yields thus a space dilation. Efficient LES scheme based on Lagrangian averaging (Meneveau and Katz 2000, and references therein) or more specifically on temporal decomposition (4) have been assessed on several prototypical flows (Chandramouli et al. 2018; Kadri Harouna and Mémin 2017; Resseguier et al. 2017d,b; Yang and Mémin 2019). Therefore, for ease of understanding, we will adhere to the vocabulary of Large Eddy Simulation (LES) and refer to the first term as the large-scale/resolved velocity component of the fluid motion while the second term will be designated as the small-scale/unresolved velocity component.

The unresolved random component is assumed to be a Gaussian field that is uncorrelated in time at the characteristic time-scale of the large-scale velocity component. Such a simple random field leads to a non-Gaussian noise when incorporated in transport equations. This stochastic formulation is related, in spirit, to the Lagrangian stochastic models based on Langevin equations proposed for turbulent dispersion (Sawford 1986) or for probability density function (PDF) modeling of turbulent flows (Haworth and Pope 1986; Pope 1994, 2000). Our interest will be, however, quite different as we only rely on such Lagrangian formulation to infer the associated large-scale Eulerian representations of the flow dynamics. In this manuscript, we will work in the Itô calculus setting. The Itô integral has the advantage of being of null expectation (it is a martingale). From time to time, it will be interesting to use the Stratonovich notation (assuming the integrand is continuous to safely move from one integral to the other) since the classical chain rule applies. Note, however, that this comes with a loss of the martingale property. As a consequence, it is interesting to highlight that the smooth in time velocity component $\boldsymbol{u}$ has the same mean as the fluid velocity $\boldsymbol{u} + \boldsymbol{\sigma}\dot{\boldsymbol{B}}_t$. A decomposition based on Stratonovich representation won't keep this property, and would separate the flow in terms of a smooth-in-time velocity component and fluctuation with non-zero mean.

The divergence-free random field involved in the Lagrangian formulation (4) is defined over the fluid domain Ω through the matrix kernel $\breve{\boldsymbol{\sigma}}(.,.,t)$ of the spatial correlation operator $\boldsymbol{\sigma}(.,t)$:

$$\forall \boldsymbol{x} \in \Omega, \quad (\boldsymbol{\sigma}(\boldsymbol{x},t)\boldsymbol{f})_i \triangleq \sum_j \int_\Omega \breve{\sigma}_{ij}(\boldsymbol{x},\boldsymbol{y},t)f_j(\boldsymbol{y},t)\mathrm{d}\boldsymbol{y}, \; i,j = 1,\ldots,d. \tag{5}$$

This operator is assumed to yield a bounded covariance tensor defined as

$$Q_{ij}(\boldsymbol{x},\boldsymbol{x}',t,t') = \mathbb{E}\big((\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t)_i(\boldsymbol{\sigma}(\boldsymbol{x}',t')\mathrm{d}\boldsymbol{B}_{t'})_j\big) = c_{ij}(\boldsymbol{x},\boldsymbol{x}',t)\delta(t-t')\mathrm{d}t,$$

with $\mathbb{E}$ denoting expectation. The last equality is obtained through Itô isometry (41). Note also that the diagonal of the covariance tensor, defined as $a_{ij}(\boldsymbol{x},t) = c_{ij}(\boldsymbol{x},\boldsymbol{x},t)$, with $c_{ij} = (\boldsymbol{\sigma}\boldsymbol{\sigma}^T)_{ij}$, corresponds to the so-called quadratic variation terms associated to the noise (c.f. Appendix A, (43), and (44)); it has the dimensions of kinematic viscosity (m^2/s) and plays the role of a *generalized* matrix-valued eddy viscosity.

3.1   Stochastic Reynolds transport theorem

The rate of change of a scalar quantity q within a volume that is transported by the random flow (4) provides us a stochastic representation of the Reynolds transport theorem (RTT) (Mémin 2014). For a divergence free random component, the stochastic RTT reads

$$\mathrm{d}\int_{V(t)} q\mathrm{d}\boldsymbol{x} = \int_{V(t)} \left(\mathrm{d}_t q + \boldsymbol{\nabla}\cdot(q\boldsymbol{u}^*)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}q - \boldsymbol{\nabla}\cdot(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}q)\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}, \tag{6}$$

where the first term represents the increment in time of the random scalar q at a fixed position, and the effective advection velocity $\boldsymbol{u}^*$ in the second term is defined as:

$$\boldsymbol{u}^* \triangleq \boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{a}. \tag{7}$$

Thus, the evolution of a conserved scalar (with an extensive property) reads immediately:

$$\mathbb{D}_t q \triangleq \mathrm{d}_t q + \boldsymbol{\nabla} \cdot (q\boldsymbol{u}^*)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \cdot \boldsymbol{\nabla}q - \boldsymbol{\nabla} \cdot (\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}q)\mathrm{d}t = 0. \tag{8}$$

For a divergence-free random field and a conserved scalar, this operator corresponds to the material derivative and the product and chain rules apply (Resseguier et al. 2017a). The resulting stochastic partial differential equation encompasses meaningful terms for studying turbulence at large scales. In fact, the fourth term on the RHS is a dissipation term (with suitable boundary conditions), since

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

with a positive semi-definite variance tensor $\boldsymbol{a}$. It depicts the mixing mechanism due to the action of the unresolved scales. The third term represents the advection of the scalar quantity by the random velocity component. This term involves the multiplication of the noise with the solution, which itself depends on the noise. Thus, it is an advective multiplicative noise and therefore non-Gaussian. The effective advection velocity in the second term captures the action of inhomogeneity of the random field on the transported scalar. It is a *statistical eddy-induced velocity* of crucial importance, as will be shown later on. Such a correction on the advection corresponds to the so-called *turbophoresis* phenomenon associated with small-scale inhomogeneity. This phenomenon drives inertial particles toward regions of lower turbulent diffusivity (Reeks 1983). In our stochastic framework, it is characterized by the turbophoresis term $\frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{a}$. It is also akin to the velocity correction introduced for tracer mean transport in oceanic or atmospheric circulation models (Andrews and McIntyre 1978). Recently, this correction was observed to play a fundamental role in the buffer layer of wall bounded flows (Pinier et al. 2019).

Incompressibility conditions for a fluid with constant density are derived from (8) as:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0, \tag{9a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u}^* = \boldsymbol{\nabla} \cdot (\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{a}) = 0. \tag{9b}$$

The first condition is intuitive and enforces a divergence free random component, whereas the second constraint imposes a divergence-free condition on the effective advection. This latter constraint provides a relation between the smooth resolved velocity component and the divergence of the variance tensor. For homogeneous random fields (such as an isotropic turbulence) this equation boils down to a classical divergence-free condition on the resolved velocity component (as the variance tensor is constant in that case). For isochoric flows with variable density, as in geophysical fluid dynamics, we get a transport equation for density of the form (8).

Moreover, given (9), the scalar transport equation has the remarkable property of energy conservation, namely,

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

where the last term on the RHS comes from the Itô integration-by-part formula (40), i.e., it conserves the main specificity of the transport equation. Thus, we refer to operator $\mathbb{D}_t$ as a stochastic transport operator. For a scalar conserved along time (such as density), this operator exactly corresponds to the material derivative (Resseguier et al. 2017a).

3.2 Stochastic Navier-Stokes equations

The stochastic RTT (6) allows us to derive from Newton's second principle (in a distributional sense) the following system of modified Navier-Stokes equations for an isochoric fluid (Mémin 2014; Resseguier et al. 2017a):

Momentum equation :

$$\mathbb{D}_t(\rho\boldsymbol{u}) = \rho\boldsymbol{g}\mathrm{d}t - \boldsymbol{\nabla}(p\mathrm{d}t + \mathrm{d}p_t^\sigma) + \mu\nabla^2(\boldsymbol{u}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t). \tag{10a}$$

Mass conservation :

$$\mathbb{D}_t\rho = 0. \tag{10b}$$

Incompressibility :

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = 0, \quad \boldsymbol{\nabla} \cdot \boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{a}) = 0. \tag{10c}$$

The forces on the RHS of (10a) are due to the gravitation potential, pressure and molecular friction forces (with the dynamic viscosity μ). The pressure term is split into a continuous pressure p and a time-uncorrelated random part $\dot{p}_t^\sigma = \mathrm{d}p_t^\sigma/\mathrm{d}t$. This latter term describes the pressure fluctuations due to the random velocity component. Note that the gravity force is continuous in time, whereas the friction force applies both on the smooth and random velocity components. For a fixed observer in a rotating frame, the rate of change of the fluid velocity incorporates the centripetal acceleration and the Coriolis acceleration as additional terms. The centrifugal force is included within an effective gravity $\boldsymbol{g}$. The Coriolis term applies both to the large-scale component of the velocity and to the random small-scale field.

This system corresponds to a large-scale description of the flow in which the effect of the random component is explicitly taken into account. Contrary to traditional Reynolds decomposition techniques, this method does not rely on the time differentiability assumption of the velocity fluctuations. The use of stochastic calculus to characterize the random component introduces naturally additional terms in the momentum equation. These terms inherently account for several interesting phenomena associated with fluid flows such as (*i*) backscattering (via the multiplicative noise), (*ii*) large-scale dissipation (through an additional diffusion term) and (*iii*) turbophoresis effect (with the advection modification). The multiplicative random term, responsible for energy backscattering, ensues from a scale separation principle between the random fluctuation and the large-scale component. In fact, if the large-scale component is assumed to be regular enough – formally a finite variation process, (10a) can be safely split in terms of martingale and finite variation terms (i.e. $\mathrm{d}t$ and $\mathrm{d}\boldsymbol{B}_t$ terms). Then, the previous system simplifies to the following stochastic LES system with a random forcing:

Momentum equations :

$$\partial_t \boldsymbol{u} + (\boldsymbol{u}^* \cdot \boldsymbol{\nabla})\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{\tau} = \boldsymbol{g} - \frac{1}{\rho}\boldsymbol{\nabla}p + \nu\nabla^2\boldsymbol{u}. \tag{11a}$$

Effective advection (with turbophoresis) :

$$\boldsymbol{u}^* = \boldsymbol{u} - \frac{1}{2}(\boldsymbol{\nabla} \cdot \boldsymbol{a}). \tag{11b}$$

Diffusion (sub-grid tensor) :

$$\boldsymbol{\tau} = (\boldsymbol{a}\boldsymbol{\nabla})\boldsymbol{u}. \tag{11c}$$

Pressure random contribution :

$$\boldsymbol{\nabla}\mathrm{d}p_t^\sigma = -\rho(\boldsymbol{\sigma}\mathrm{d}\hat{\boldsymbol{B}}_t \cdot \boldsymbol{\nabla})\boldsymbol{u} + \nu\nabla^2\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t. \tag{11d}$$

Mass conservation :

$$\mathbb{D}_t\rho = 0. \tag{11e}$$

Here, the momentum equation is a classical PDE (where the kinematic viscosity $\nu = \mu/\rho$ has been introduced) with a random forcing through the density. The mass conservation (11e) remains a stochastic PDE. For the particular case of constant density, this equation simplifies to the following incompressibility constraints,

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = 0, \quad \boldsymbol{\nabla} \cdot \boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{a}) = 0. \tag{12a}$$

In this case, the momentum equation is purely deterministic (without random forcing). Nevertheless, the system incorporates a stochastic balance equation on the pressure contribution $\mathrm{d}p_t^\sigma$ associated to the unresolved random term. Also, this system includes the effective advection and the large-scale diffusion brought by the random component, but loses in the momentum equation the multiplicative random term responsible for energy backscattering. The shape of the dissipation operator is a priori known. Thus, there is no need to invoke the Boussinesq assumption to model the Reynolds tensor. Note that the variance tensor (and the noise, if required) can be constructed by drawing inspiration from known LES sub-grid scale models such as the Smagorinsky model or its dynamical version based on Germano decomposition(Chandramouli et al. 2018; Kadri Harouna and Mémin 2017). However, it was concluded that local empirical constructions of Chandramouli et al. (2018); Kadri Harouna and Mémin (2017) for the variance tensor provided more efficient LES schemes. Oceanic subgrid scale models such as the Gent-McWilliams model can also be constructed from the LU framework as shown in (Mémin 2014).

3.3 Stochastic geophysical models

In order to account for stratified oceanic or atmospheric flows, the pressure and density are decomposed as

$$p = \widetilde{p}(z) + p'(x, y, z, t), \tag{13}$$

$$\rho = \widetilde{\rho}(z) + \rho'(x, y, z, t). \tag{14}$$

The scalar fields $\widetilde{\rho}(z) \triangleq \rho_b + \rho_0(z)$ and $\widetilde{p}(z)$ are stationary components in equilibrium that depend only on height. The fluctuating components are random functions, i.e. they depend on the random velocity component. From the random momentum equation (10a), it can be readily inferred that the stationary components are related by the hydrostatic balance:

$$\frac{\partial \widetilde{p}}{\partial z} = -g(\rho_b + \rho_0(z)), \tag{15}$$

where g denotes the acceleration due to gravity. Assuming small density fluctuations (i.e the so-called Boussinesq approximation), the density anomalies are stochastically transported by

$$\mathbb{D}_t(\rho - \rho_b) = 0. \tag{16}$$

Including the Coriolis correction force due to earth rotation, $f\boldsymbol{k}$, we obtain the so-called simple Boussinesq system of equations:

Momentum equations:

$$\mathbb{D}_t\boldsymbol{u} + f\boldsymbol{k} \times (\boldsymbol{u}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = b\,\boldsymbol{k}\mathrm{d}t - \frac{1}{\rho_b}\boldsymbol{\nabla}(p'\mathrm{d}t + \mathrm{d}p_t^\sigma) + \mathscr{F}(\boldsymbol{u}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t). \tag{17a}$$

Effective advection:

$$\boldsymbol{u}^* = \begin{pmatrix} \boldsymbol{u}_h^* \\ w^* \end{pmatrix} = \boldsymbol{u} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T. \tag{17b}$$

Buoyancy equation:

$$\mathbb{D}_t b + N^2\left(w^*\mathrm{d}t + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\right) = \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}_{\bullet z}N^2\right)\mathrm{d}t, \tag{17c}$$

$$b = -\frac{g}{\rho_b}\rho', \quad N^2(z) = -\frac{g}{\rho_b}\partial_z\rho_0(z). \tag{17d}$$

Incompressibility:

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}^* = \boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = 0. \tag{17e}$$

The term $\mathscr{F}(\boldsymbol{u})$ stands for the friction force. Neglecting this term yields the Euler-Boussinesq model. As is common in geophysical models, the velocity is split into horizontal and vertical components. As a consequence, the effective drift is also split into the corresponding horizontal $\boldsymbol{u}_h^*$ and vertical w^* components. The dynamics of density fluctuations is expressed through the buoyancy variable $b = -\frac{g}{\rho_b}\rho'$ and the Brunt-Väisälä stratification frequency $N^2(z) = -\frac{g}{\rho_b}\partial_z\rho_0(z)$. Also, $\boldsymbol{a}_{\bullet z}$ stands for the z column vector of $\boldsymbol{a}$. In comparison to the previous stochastic Navier-Stokes system, the buoyancy term in the momentum equation constitutes an additional random forcing of the vertical large-scale velocity component. As previously shown, this system can also be turned into an LES system with a scale separation assumption (i.e. finite variation of the large-scale component). However, in this case the momentum equation retains its random nature due to the buoyancy forcing.

From this system of equations a diverse set of approximated models can be obtained through adimensionalization and power series expansions in terms of small Rossby number (ratio of the inertial force to Coriolis force) with proper scalings (see e.g. Appendix D). These developments follow the same path as in the deterministic setting. However, the noise introduces here an additional degree of freedom that must be appropriately accounted for (Resseguier et al. 2017a,b,c). Instances of these approximated models include the Planetary Geostrophic model (PG), the Quasi-Geostrophic model (QG) and the Surface Quasi-Geostrophic model (SQG).

Stochastic reduced order models obtained from Galerkin projection onto data-driven basis functions can also be derived from such a system of equations when combined with snapshots of velocity data (Resseguier et al. 2017d). The stochastic Boussinesq model for Rayleigh-Benard convection leads to a stochastic version of the Lorenz system (Chapron et al. 2018). This system, used as a proof of concept, seems to show an ability of the stochastic system to visit the attractor region of the deterministic system more efficiently than the deterministic model even for noise of low amplitude (Chapron et al. 2018).

In the following section we focus on the effective advection and on the interactions it has with the large-scale components. This term unveils the contribution of inhomogeneity at the small-scales as a driver to trigger large-scale structures in the flow.

4 Stochastic momentum equations in terms of effective advection

In order to interpret the action of the turbophoresis term in (17b) on the large-scale flow, we rewrite the momentum equation in an equivalent form outlining the contribution of the effective advection. For reasons that will be made clear in the following, the turbophoresis term $\boldsymbol{u}_S \stackrel{\triangle}{=} \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}$ is, henceforth, referred to as the "Itô-Stokes" drift. In its expanded form, the stochastic Euler momentum equation (17a) reads:

$$\mathrm{d}_t\boldsymbol{u} + \left((\boldsymbol{u}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\right)\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{u})\mathrm{d}t + \boldsymbol{f}\times(\boldsymbol{u}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = b\,\boldsymbol{k}\mathrm{d}t - \boldsymbol{\nabla}(p'\mathrm{d}t + \mathrm{d}p_t^\sigma). \tag{18}$$

By applying a change of variable in (18) from $\boldsymbol{u}$ to $\boldsymbol{u}^*$, assuming a stationary Itô-Stokes drift, we get,

$$
\begin{aligned}
\mathrm{d}_t\boldsymbol{u}^* + \bigg(\big((\overbrace{\boldsymbol{u}^* - \boldsymbol{u}_S}^{\text{turbophoresis}})\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \big) \cdot \boldsymbol{\nabla} \bigg)\boldsymbol{u}^* &- \overbrace{\frac{1}{2}\boldsymbol{\nabla}\cdot\big((\boldsymbol{a}\boldsymbol{\nabla})(\boldsymbol{u}^*+\boldsymbol{u}_S)\big)}^{SGS}\mathrm{d}t - \overbrace{(\boldsymbol{u}_S\cdot\boldsymbol{\nabla})\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}^{\substack{\text{noise advection}\\\text{by Itô-Stokes}}} = b\,\boldsymbol{k}\mathrm{d}t \\
- \boldsymbol{\nabla}\mathrm{d}\pi_t &- \underbrace{\boldsymbol{\omega}_S\times(\boldsymbol{u}^*\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{Itô-Stokes force}} - \boldsymbol{f}\times(\boldsymbol{u}^*\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) - \underbrace{\boldsymbol{f}\times\boldsymbol{u}_S}_{\text{Coriolis Stokes}}\,\mathrm{d}t + \underbrace{\boldsymbol{u}_S\times\big(\boldsymbol{\omega}^*\mathrm{d}t+\boldsymbol{\nabla}\times(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\big)}_{\text{vortex force}},
\end{aligned}
\tag{19}
$$

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

$$
\mathrm{d}\pi_t = \frac{1}{\rho_b}(p'\mathrm{d}t+\mathrm{d}p_t^\sigma) + \frac{1}{2}\big((\|\boldsymbol{u}^*\|^2+\|\boldsymbol{u}_S\|^2)-\|\boldsymbol{u}^*-\boldsymbol{u}_S\|^2\big)\mathrm{d}t + \boldsymbol{u}_S\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t,
\tag{20}
$$

by using the vector identity $\boldsymbol{\nabla}(\boldsymbol{u}\cdot\boldsymbol{v}) = (\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{v}+(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{u}+\boldsymbol{u}\times\boldsymbol{\nabla}\times\boldsymbol{v}+\boldsymbol{v}\times\boldsymbol{\nabla}\times\boldsymbol{u}$. On the RHS of equation (19), a Craik-Leibovich vortex force appears. This force depicts the statistical contribution of the inhomogeneity carried by the variance of the random field on the large-scale current. Hence, this momentum equation may be seen as a general stochastic expression of the Craik-Leibovich system in which the turbophoresis term replaces the Stokes drift associated to wave motion. This is the reason why we designate this term with the more general descriptive name of Itô-Stokes drift. This equation still includes a turbophoresis term and a large-scale dissipation. The Coriolis force now includes a correction term that depends on the Itô-Stokes drift. An advection of the small-scale component by the Itô-Stokes drift also emerges. In addition to the vortex force, this stochastic formulation includes another force, referred to here as the *Itô-Stokes force*, related to the interaction of the flow with the vorticity of the Itô-Stokes drift. This force could also be gathered with the Coriolis term to form a Coriolis force corrected by the Itô-Stokes vorticity. This allows us to notice an interesting particular case. Given a downward curl of the Itô-Stokes drift and a negligible Coriolis force in front of the Itô-Stokes force, a correction term appears acting opposite to the Coriolis force. This is in agreement with recent studies showing the variability in turbulent intensity to wind direction and latitude (Glazunov 2010; Liu et al. 2018). Note that the Itô-Stokes drift depends on the variance tensor $\boldsymbol{a}$, which is defined as the variance of the random displacement $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)$ divided by a decorrelation time τ (i.e. $a\propto L^2/\tau$). Thus, the divergence and curl of the Itô-Stokes drift both scale as the inverse of this decorrelation time (i.e. $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot a\propto 1/\tau$). These properties, possibly leading to Itô-Stokes drift pumping effects, are thus small for sufficiently large decorrelation times such as infra gravity or long waves. For gravity waves, the divergence and curl of the Itô-Stokes drift may not be necessarily small, only becoming negligible for small slope waves and spatially smooth waves of low amplitude. For isotropic random fluctuations, as in isotropic turbulence, the variance tensor is constant in space and the Itô-Stokes drift cancels. In that case, the effective advection and the large-scale velocity are identical and there is no vortex force nor Itô-Stokes force to structure the large-scale flow components. In such a scenario, the only interaction between the large-scale velocity component and the fluctuations is the large-scale diffusion term and the random advection term. Both of them are in equilibrium to ensure energy conservation. The ratio between the decorrelation time and the characteristic time of the resolved quantities (or equivalently the ratio between the Itô-Stokes drift divergence and the characteristic time at the resolved scale) can be expressed as a Reynolds number based on the variance tensor (or eddy viscosity a) : $\frac{\tau}{T}=\frac{L^2}{a}\frac{u}{L}=Re(a)$. This Reynolds number corresponds to a measure of the decorrelation time with respect to the the resolved scale characteristic time. From the previous discussion, it is also a measure of the Itô-Stokes drift divergence associated with the so-called Stokes-drift pumps responsible for upwelling and downwelling occurring at the rear and in front of groups of surface gravity waves respectively (Haney and Young 2017; Longuet-Higgins and Stewart 1964).

In the following we show how equation (19) can be simplified to get the Craik-Leibovich equation.

4.1 From LU equations to Craik-Leibovich equations

Equation (19) can be simplified with additional assumptions on the inhomogeneity of the small-scale component. Given the incompressibility condition on the effective advection, $\boldsymbol{\nabla}\cdot\boldsymbol{u}^*=\boldsymbol{\nabla}\cdot(\boldsymbol{u}-\boldsymbol{u}_S)=0$, and an incompressible large-scale flow, we infer an incompressible Itô-Stokes drift $\boldsymbol{\nabla}\cdot\boldsymbol{u}_S=0$. In this case, the summation of the SGS contribution on $\boldsymbol{u}^*$ and the turbophoresis term gives:

$$
-\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{u}^*)-(\boldsymbol{u}_S\cdot\boldsymbol{\nabla})\boldsymbol{u}^* = -\frac{1}{2}\sum_{i,j=1}^d \partial_{x_i}\partial_{x_j}(a_{ij}\boldsymbol{u}^*),
\tag{21}
$$

and the Itô-Stokes diffusion term reads:

$$
-\frac{1}{2}\boldsymbol{\nabla}\cdot\big((\boldsymbol{a}\boldsymbol{\nabla})\boldsymbol{u}_S\big) = -\frac{1}{2}\sum_{i,j=1}^d \partial_{x_i}\partial_{x_j}\big(a_{ij}\boldsymbol{u}_S\big)+(\boldsymbol{u}_S\cdot\boldsymbol{\nabla})\boldsymbol{u}_S.
\tag{22}
$$

The last term of this equation can be written:

$$\left(\boldsymbol{u}_S \cdot \boldsymbol{\nabla}\right)\boldsymbol{u}_S = \frac{1}{2}\boldsymbol{\nabla}(\|\boldsymbol{u}_S\|^2) - \boldsymbol{u}_S \times \boldsymbol{\omega}_S. \tag{23}$$

Equation (19) can thus be written as

$$\mathrm{d}_t\boldsymbol{u}^* + \left((\boldsymbol{u}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla}\right)\boldsymbol{u}^* - \overbrace{\frac{1}{2}\sum_{i,j=1}^{d}\partial_{x_i}\partial_{x_j}\left(a_{ij}(\boldsymbol{u}^* + \boldsymbol{u}_S)\right)\mathrm{d}t}^{SGS} - \overbrace{\left(\boldsymbol{u}_S \cdot \boldsymbol{\nabla}\right)\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}^{\text{noise advection}} =$$
$$b\,\boldsymbol{k}\mathrm{d}t - \boldsymbol{\nabla}(\mathrm{d}\pi_t + \frac{1}{2}\|\boldsymbol{u}_S\|^2\mathrm{d}t) - \underbrace{\boldsymbol{\omega}_S \times \left((\boldsymbol{u}^* + \boldsymbol{u}_S)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)}_{\text{Itô-Stokes force}} - \boldsymbol{f} \times (\boldsymbol{u}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) -$$
$$\underbrace{\boldsymbol{f} \times \boldsymbol{u}_S}_{\text{Coriolis Stokes}}\ \mathrm{d}t + \underbrace{\boldsymbol{u}_S \times \left(\boldsymbol{\omega}^*\mathrm{d}t + \boldsymbol{\nabla} \times (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\right)}_{\text{vortex force}}. \tag{24}$$

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

$$\mathrm{d}_t\boldsymbol{u}^* + \left((\boldsymbol{u}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla}\right)\boldsymbol{u}^* - \overbrace{\frac{1}{2}\sum_{i,j=1}^{d}\partial_{x_i}\partial_{x_j}\left(a_{ij}(\boldsymbol{u}^* + \boldsymbol{u}_S)\right)\mathrm{d}t}^{SGS} - \overbrace{\left(\boldsymbol{u}_S \cdot \boldsymbol{\nabla}\right)\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}^{\text{noise advection}} = b\,\boldsymbol{k}\mathrm{d}t$$
$$- \boldsymbol{\nabla}(\mathrm{d}\pi_t + \frac{1}{2}\|\boldsymbol{u}_S\|^2\mathrm{d}t) - \underbrace{(\boldsymbol{\omega}_S + \boldsymbol{f}) \times \left((\boldsymbol{u}^* + \boldsymbol{u}_S)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)}_{\text{Coriolis + Itô-Stokes force}} + \underbrace{\boldsymbol{u}_S \times \left(\boldsymbol{\omega}^*\mathrm{d}t + \boldsymbol{\nabla} \times (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\right)}_{\text{vortex force}}. \tag{25}$$

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

$$\mathrm{d}_t\boldsymbol{u}^* + \left((\boldsymbol{u}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla}\right)\boldsymbol{u}^* - \overbrace{\frac{1}{2}\sum_{i,j=1}^{d}\partial_{x_i}\partial_{x_j}\left(a_{ij}(\boldsymbol{u}^* + \boldsymbol{u}_S)\right)\mathrm{d}t}^{SGS} - \overbrace{\left(\boldsymbol{u}_S \cdot \boldsymbol{\nabla}\right)\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}^{\text{noise advection}} =$$
$$b\boldsymbol{k}\,\mathrm{d}t - \boldsymbol{\nabla}(\mathrm{d}\pi_t + \frac{1}{2}\|\boldsymbol{u}_S\|^2\mathrm{d}t) - \underbrace{\boldsymbol{f} \times \left((\boldsymbol{u}^* + \boldsymbol{u}_S)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)}_{\text{Coriolis + Stokes}}\mathrm{d}t + \underbrace{\boldsymbol{u}_S \times \left(\boldsymbol{\omega}^*\mathrm{d}t + \boldsymbol{\nabla} \times (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\right)}_{\text{vortex force}}. \tag{26}$$

As previously described, these equations can be written in an LES deterministic form (assuming $\boldsymbol{u}^*$ is a finite variation process) through a decomposition in terms of martingale and finite variation terms. For the large-scale velocity component this yields (from (25)) a momentum equation of the form,

$$\partial_t\boldsymbol{u}^* + (\boldsymbol{u}^* \cdot \boldsymbol{\nabla})\boldsymbol{u}^* - \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \cdot \left(a(\boldsymbol{u}^* + \boldsymbol{u}_S)^T\right) = b\,\boldsymbol{k} - \boldsymbol{\nabla}(\pi + \frac{1}{2}\|\boldsymbol{u}_S\|^2) - \underbrace{\boldsymbol{f} \times (\boldsymbol{u}^* + \boldsymbol{u}_S)}_{\text{Coriolis + Stokes}} + \underbrace{\boldsymbol{u}_S \times \boldsymbol{\omega}^*}_{\text{vortex force}}, \tag{27}$$

which recovers the CL form (1) with a pressure term

$$\pi = p' + \frac{1}{2}\|\boldsymbol{u}^* + \boldsymbol{u}_S\|^2 - \frac{1}{2}\|\boldsymbol{u}^*\|^2 = p' + \frac{1}{2}\|\boldsymbol{u}\|^2 - \frac{1}{2}\|\boldsymbol{u}^*\|^2, \tag{28}$$

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

In the original CL theory, the Stokes drift is identified with the residual velocity of the fast orbital motion of the waves. It arises from a linear perturbation theory and can be defined (in the assumption of small surface wave slope and nearly irrotational wave motion) as the difference between the Lagrangian flow velocity of a fluid element and the averaged Eulerian flow velocity at a fixed point. In our stochastic framework, it corresponds, instead, to a statistical correction resulting from inhomogeneity of the small-scale. The quasi-harmonic assumption used to get the original CL system is much stronger than the usual small slope wave assumption supporting the CL derivation (Leibovich 1980).

As discussed previously, this quasi-harmonic assumption considers either a sufficiently large decorrelation time for the variance of the velocity fluctuations or smooth enough velocity fields with small inhomogeneity. Within these hypotheses pertaining to the derivation of the Craik-Leibovich equations from the LU framework (i.e deterministic $\boldsymbol{u}^*$ and stationary quasi-harmonic Itô-Stokes drift), the effective advection $\boldsymbol{u}^*$ can be associated to the mean Eulerian component whereas $\boldsymbol{u}$ is, by definition, a smooth Lagrangian velocity component. Without these assumptions, we get the more generalized system (19) in which the Itô-Stokes drift component interacts with all the velocity components and with itself through a Lamb vector involving its curl. It is important to stress that this general stochastic CL equation is equivalent to the Euler momentum equations (18) for it preserves kinetic energy. This link between the stochastic Euler equations and the Craik-Leibovich equations clearly shows the potential impact of inhomogeneity of the small-scale random field in shaping large-scale structures such as Langmuir cells. Let us remark that for the buoyancy equation (17c) or the transport of any scalar no particular changes needs to be done in our setting. These transport equations involve the effective advection $\boldsymbol{u}^*$ which includes the Itô-Stokes drift correction.

4.2 Practical general consequences

The principal consequence of the connection between the Craik-Leibovich system and the stochastic system (17a) is that in the latter the explicit inclusion of the vortex force is not necessary to trigger secondary circulations. In the stochastic setting, such circulations require only a small-scale velocity component with an appropriately defined Itô-Stokes drift. This constitutes a simplified procedure than considering an LES representation with an explicit vortex force and the accompanying associated modifications (Coriolis modification; fluctuation-fluctuation interaction; modified pressure, etc.).

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

Before examining some numerical results, we provide some general remarks on the LU framework and on the associated choices.

4.3 General remarks on the LU framework

Here, we address several points, which, in our opinion, characterizes our LU framework:

- The transport equations derived under the LU framework involves an effective advection $\boldsymbol{u}^*$ independent of the stochastic integral applied (i.e. both in Stratonovich and Itô setting) (see (51) and (59) respectively). The LU framework under either stochastic setting retains the Itô-Stokes drift associated to the inhomogeneity of the fluctuations, which is, as shown in this work, an important driver of secondary circulations. **Hence, this capability is an intrinsic property of the LU framework in the sense that it is independent of the type of stochastic integral used**.

- The LU framework based on decomposition (4) includes a centered noise (i.e. of zero-mean) by the definition of an Itô integral. Therefore, the mean of the Lagrangian large-scale velocity component corresponds to the mean of the Lagrangian fluid velocity. In contrast, a Stratonovich decomposition would lead to non-zero mean for the fluctuations.

  For instance, let us consider the following decomposition :

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}'(\boldsymbol{X}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t) \circ \mathrm{d}\boldsymbol{B}_t. \tag{29}$$

  Such a Stratonovich representation with divergence-free constraint is equivalent, as shown in Appendix C, to the following Itô form :

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}'(\boldsymbol{X}_t, t)\mathrm{d}t + \frac{1}{2}\boldsymbol{\nabla} \cdot \Big(\boldsymbol{a}(\boldsymbol{X}_t, t)\Big)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\mathrm{d}\boldsymbol{B}_t.$$

  A comparison of this equation with (4) and (7) gives $\boldsymbol{u}' = \boldsymbol{u}^* = \boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{a}$. As a consequence, working directly with a Stratonovich decomposition, for instance as done by (Holm 2015), leads to a momentum equations on $\boldsymbol{u}^*$ in which there is no effective advection. In this case, the relation with Craik-Leibovich system can only be obtained by switching to the Itô setting or by considering additional random processes c.f. (Holm 2019). Note also that a Stratonovich decomposition requires stronger smoothness assumptions both in space and time of the transported quantities (see theorem C.2).

- The LU framework conserves the energy of a transported scalar for every realization. As shown later for QG models, this allows us to build dynamics and draw realizations that have exactly the same energy conservation properties as their deterministic counterpart. This strong asset of the LU framework enables us to propose efficient schemes for numerical simulation, analysis and data assimilation of 3D turbulent flows (Chandramouli et al. 2018; Kadri Harouna and Mémin 2017; Resseguier et al. 2017d).

- The derivation of the stochastic models under the LU framework follows the same path as the deterministic derivation. As shown in this work, only the noise and its amplitude needs to be properly scaled. To oceanographers, this may provide a very interesting and practical tool for investigating the implication of the small-scales.

In the following, we apply our LU framework to a barotropic quasi–geostrophic model (BQG). We show that the introduction of inhomogeneous noise induces a structuration of the flow with strong secondary vortices.

5 Stochastic Barotropic Quasi–Geostrophic model

The derivation of a BQG system under location uncertainty (denoted as BQG$_{\mathrm{LU}}$), described in Appendix D, follows the same strategy as in the classical framework.

The stochastic advection–diffusion equation of the potential vorticity (PV) q reads

$$\mathbb{D}_t q = \mathrm{d}R_t, \tag{30a}$$

$$q = \nabla^2 \psi + f - \frac{1}{L_R^2}\psi, \tag{30b}$$

$$\boldsymbol{u} = \boldsymbol{\nabla}^\perp \psi, \quad \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp \varphi \mathrm{d}B_t, \quad \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot\boldsymbol{a} = 0, \tag{30c}$$

where $f = f_0 + \beta y$ is the Coriolis parameter on a beta-plane, L_R is the Rossby deformation radius, $\boldsymbol{\nabla}^\perp = [-\partial_y, \partial_x]^T$ is the perpendicular gradient in two dimensions, $\varphi \mathrm{d}B_t$ is a random streamfunction for the unresolved flow and the source/sink process $\mathrm{d}R_t$ is given by

$$\mathrm{d}R_t \triangleq \Big[\frac{1}{2}\mathrm{tr}\big(\boldsymbol{\nabla}^\perp(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\boldsymbol{\nabla}\boldsymbol{u}^T\big) + \frac{1}{2}\boldsymbol{\nabla}\cdot\big(\frac{\partial\boldsymbol{a}}{\partial x}\boldsymbol{\nabla}v - \frac{\partial\boldsymbol{a}}{\partial y}\boldsymbol{\nabla}u\big) - (\boldsymbol{\nabla}\cdot\boldsymbol{a})\cdot\boldsymbol{\nabla}f\Big]\mathrm{d}t -$$
$$\mathrm{tr}\big(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T\boldsymbol{\nabla}\boldsymbol{u}^T\big). \tag{30d}$$

To interpret these forcing terms, it is informative to write the PV evolution equation in Stratonovich notation:

$$\mathbb{D}_t^S q = \mathrm{d}R_t^S, \tag{31a}$$

$$\mathrm{d}R_t^S \triangleq \frac{1}{2}\Big[\mathrm{tr}\big(\boldsymbol{\nabla}^\perp(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\boldsymbol{\nabla}\boldsymbol{u}^T\big) - (\boldsymbol{\nabla}\cdot\boldsymbol{a})\cdot\boldsymbol{\nabla}f\Big]\mathrm{d}t - \mathrm{tr}\big(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)^T\boldsymbol{\nabla}\boldsymbol{u}^T\big). \tag{31b}$$

The first and last terms can now be interpreted as the rotating interaction between the strain vectors of the resolved and unresolved flows. For instance, $\mathrm{tr}\Big(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big) = \frac{1}{2}\det\Big(\mathrm{S}(\varphi\circ\mathrm{d}\boldsymbol{B}_t), \mathrm{S}(\psi)\Big)$ with S the strain vector (McWilliams 1984; Weiss 1991) can be decomposed into $\mathrm{S}_1(\psi) = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = -2\frac{\partial^2\psi}{\partial x\partial y}$ and $\mathrm{S}_2(\psi) = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2\psi}{\partial y^2}$. In particular, this anti-symmetric source term will be zero when the two strain vectors are colinear. A more precise exploration about these terms is described in (Resseguier et al. 2017b). In the following, we outline their contributions to the conservative budget of the QG system.

System (30) has the remarkable property of (resolved) total energy E conservation along each realization, for ideal periodic, free or no-slip boundary conditions:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = 0, \quad E = \int_\Omega \frac{1}{2}\|\boldsymbol{\nabla}\psi\|^2 + \frac{1}{2}\Big(\frac{\psi}{L_R}\Big)^2. \tag{32}$$

This is shown in Appendix E using the Stratonovich notation (76a). The energy of the source processes $\mathrm{d}R_t^S$ compensates the increase of energy due to the advection of the resolved flow by the unresolved one. Therefore, the stochastic BQG system keeps an essential characteristic of the BQG deterministic system, though, in this setting, PV is not strictly conserved. We note also that this system introduces a strong incompressiblility constraint on the Itô-Stokes drift (30b).

5.1 Numerical experiments

In this section we aim at comparing, for BQG$_{\mathrm{LU}}$, the effect of an isotropic homogeneous noise with a specified spectrum (and a null Itô-Stokes drift) and an inhomogeneous noise built from a scale similarity assumption. For the first noise, no large-scale secondary structuration should be observed as no Itô-Stokes drift is associated whereas in the second case the Itô-Stokes drift is non-zero and should impact the large-scale solution. The construction of these two noises are presented hereafter.

5.1.1 The homogeneous isotropic stationary model

In the context of a stochastic SQG simulation, (Resseguier et al. 2017b) elaborated a random streamfunction $\varphi\mathrm{d}B_t$ defined through a streamfunction kernel $\breve{\varphi}$:

$$\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp\varphi(\boldsymbol{x})\mathrm{d}B_t = (\boldsymbol{\nabla}\breve{\varphi}\star\mathrm{d}B_t)(\boldsymbol{x}). \tag{33}$$

In practice, the streamfunction kernel is specified by three parameters: a fixed omni-directional spectrum slope, denoted s; a band-pass filter f_{BP} with support in the range of two wavenumbers κ_m and κ_M, and a scalar variance tensor a_0 similar to an eddy viscosity coefficient. In Fourier space the random streamfunction is then conveniently defined as:

$$\widehat{\varphi(\boldsymbol{x})\mathrm{d}B_t}(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}} f_{BP}(\|\boldsymbol{k}\|)\|\boldsymbol{k}\|^{-\alpha}\widehat{\eta}_t(\boldsymbol{k}) \text{ with } \alpha = (3+s)/2, \tag{34}$$

where η_t is a space-time white noise process and A is a constant to ensure $\mathbb{E}\|\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t\|^2 = 2a_0\Delta t$.

5.1.2 The heterogeneous non–stationary model

It is possible to write the unresolved velocity through a Karhunen-Loeve decomposition (Memin 2014):

$$\boldsymbol{\sigma}(\boldsymbol{x},t)\dot{\boldsymbol{B}}_t = \sum_{k=1}^N \boldsymbol{\xi}_k(\boldsymbol{x},t)\dot{W}_k, \tag{35}$$

where $(\boldsymbol{\xi}_k)_{k=1,\dots,N}$ are the so-called Empirical Orthogonal Functions (EOFs) weighted by the corresponding eigenvalues, W_k are independent one-dimensional Brownian motions and N is possibly infinite. That kernel corresponds – up to a decorrelation time – to the one-time-two-points covariance of the unresolved velocity.

Accordingly, we propose here a new parameterization approach with time-dependent EOFs estimated on-line from a coarse simulation. To that end, assuming a scale similarity assumption, a set of small-scale pseudo-observations are generated directly from the resolved velocity field. Then, a singular value decomposition (SVD) of the associated empirical covariance matrix provides the EOF representation. The variance of the noise is scaled at the resolution scale using a turbulence power-law scaling (Kadri Harouna and Mémin 2017).

More precisely, at the current simulation time t, n_O pseudo-observations of the velocity fluctuations $\boldsymbol{v}'$ at each grid point are build from a $n_w \times n_w$ (n_w odd) sliding window. The fluctuations for a given realization and a given grid point are set from the anomaly between the empirical local velocity average in the spatially centered window and the velocity associated to random draws within this local window. We denote by $L = n_w \ell$ the spatial scale of the window, where ℓ is the smallest scale of the simulation. A SVD of the empirical correlation matrix composed of these pseudo-observations is performed to obtain the EOFs, defined as in (35) with $N = n_o$. The pseudo-realizations $\boldsymbol{v}' = \boldsymbol{v}'_L$ have been generated through a $n_w \times n_w$ window average and are defined at a resolution corresponding to a virtual observation scale L; they must be scaled down to the "simulation scale" ℓ. Following (Kadri Harouna and Mémin 2017) the unresolved velocity variance tensor $\boldsymbol{a}$ is rescaled by a coefficient:

$$\boldsymbol{a}_\ell = n_w^{-2/3} \boldsymbol{a}_L, \tag{36}$$

where $\boldsymbol{a}_L$ and $\boldsymbol{a}_\ell$ are the variance tensors at scales L and ℓ, respectively. The unresolved velocity is finally defined as:

$$\boldsymbol{\sigma}^{(\ell)} \dot{\boldsymbol{B}}_t = n_w^{-1/3} \boldsymbol{\sigma}^{(L)} \dot{\boldsymbol{B}}_t. \tag{37}$$

5.1.3 Simulation results

The non-dimensionalized simulation of the $\mathrm{BQG_{LU}}$ system (with $\beta = L_R = 1$) for the two types of noise are initialized with the same monochromatic Rossby waves, $\psi(\boldsymbol{x}, t = 0) = 0.1 \cos(2x)$, (which is a solution of the deterministic BQG system). The geometry is defined as a uniform 128×128 Cartesian grid within a doubly periodic domain $[0, 2\pi] \times [0, 2\pi]$. Hence, all prognostic variables, such as the streamfunction ψ, the relative vorticity $\xi = \nabla^2 \psi$, the vector noise $\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t$ and the variance tensor $\boldsymbol{a}$ are double periodic. To discretize spatially the vorticity equation (30a), we employ Arakawa's nine-point conservative scheme (Arakawa and Lamb 1977), together with second order centered finite differences for the stochastic and diffusion terms. For time-stepping of the large-scale time-correlated terms in (30a), we use a strong stability preserving third order Runge Kutta (RK3) scheme (Gottlieb 2005) with a Courant-Friedrich-Lewy (CFL) number of 1/3. The time-uncorrelated terms in (30a) are integrated in the final step of the RK3 scheme using the Euler-Maruyama scheme (Gugole and Franzke 2019; Pavliotis and Stuart 2008). To invert the Helmholtz equation associated to the streamfunction (30b), an efficient Fast Fourier Transform (FFT) solver (Press et al. 2007) is adopted.

The objective here, in addition to the study of the structuration effect by the small-scales, is to assess the preservation of the large-scale initial Rossby wave for the stochastic system under both noise models. In the inhomogeneous case, at each time we uniformly draw $n_o = 20$ pseudo-observations by sliding a $n_w^2 = 3 \times 3$ window over the grid, with circular boundary conditions. Both the noise models are fixed to a similar amplitude. To that end, the homogeneous noise amplitude is fixed with $a_0 \approx 2.85 \times 10^{-5}$, which is determined by the mean amplitude of the inhomogeneous (and time varying) noise. The other parameters are chosen as: $k_M = \pi / \Delta$ with Δ the grid spacing, $k_m = k_M / 2$ and $s = -3$. The vorticity evolution for each noise is plotted in figures 1-2.

As can be observed, both noises lead to stable solutions. The initial monochromatic wave can still be seen in both simulations in figures 1 and 2. The monochromatic wave can also be seen in the Fourier spectrum of the temporal signal associated to a given point of the grid (see Fig. 3). The deterministic BQG and the $\mathrm{BQG_{LU}}$ for both noises have a strong common frequency peak corresponding to this monochromatic wave. In the homogeneous case, for which the Itô-Stokes drift is null, we observe a statistically homogeneous solution with no particular structuration of the flow. This can be observed visually in Fig. 1 or infered from the Fourier spectra in Fig. 3. We clearly see the superposition of the monochromatic wave with a homogeneous noise pattern active at all scales. In contrast, for the inhomogeneous case, after a spin-up time of approximately $t = 200$, the apparition of large vortices can be observed in Fig. 2. These patterns correspond to the secondary peak on the energy spectrum at slightly lower frequency than the monochromatic initial wave as observed in the RHS of Fig. 3. They correspond to slower events. These vortices remain stable for a while then disappear and reappear with a longer time periodicity. This structuration is stable along time and still conserves the large scale Rossby waves. Note that due to energy conservation, a part of the energy of the monochromatic wave is redistributed to the secondary vortex structure (as see in Fig. 3).
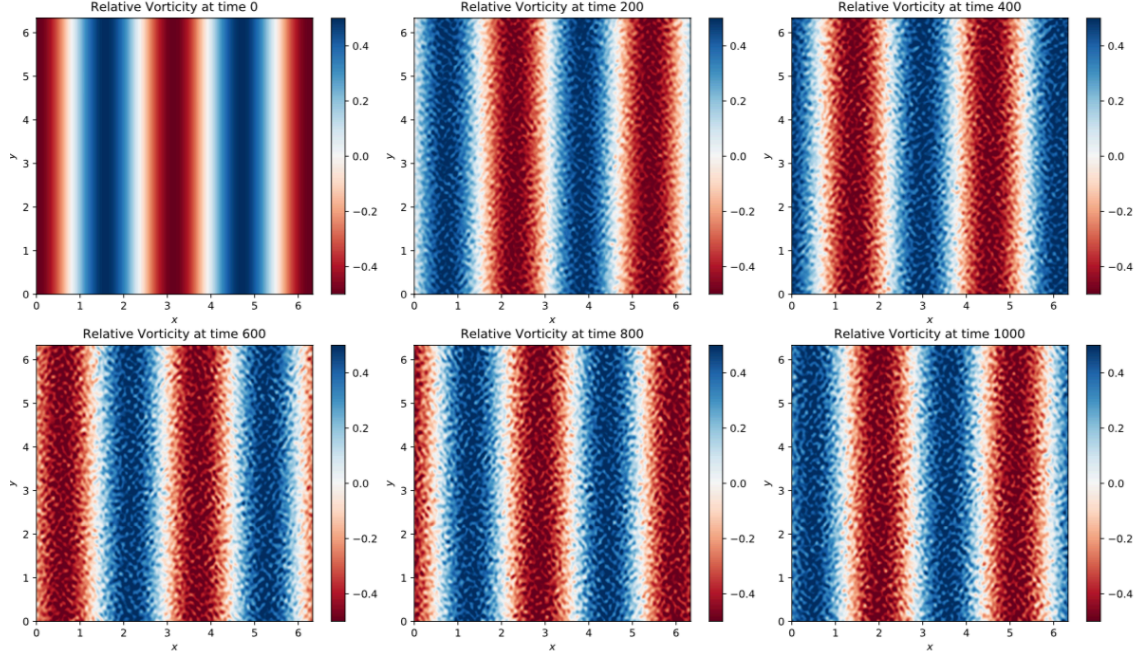
Figure 1: BQG$_{LU}$ evolution of the relative vorticity for the isotropic stationary noise model (section 5.1.a) over a period of $t = 1000$ adimensional time and a 128×128 spatial grid. At the initial time (left figure in the first row) the relative vorticity is initialized with a monochromatic wave (solution of the deterministic BQG model). This monochromatic wave is conserved by the stochastic system and an homogeneous noisy component superimposed on it.
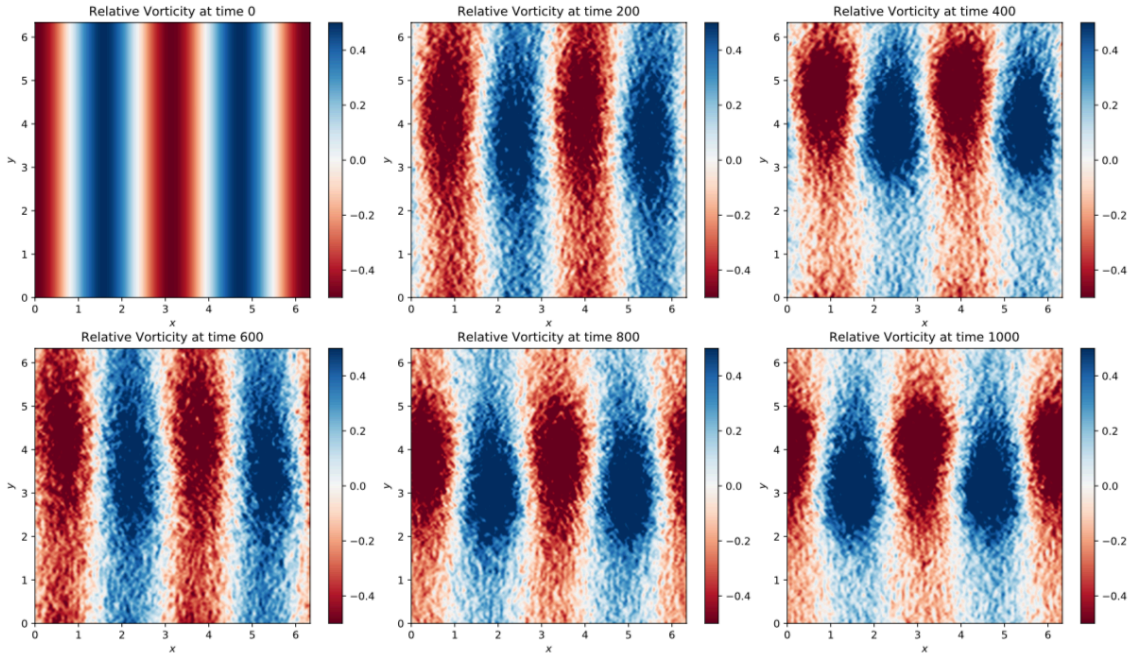
Figure 2: BQG$_{LU}$ evolution of the relative vorticity for the heterogeneous non–stationary noise model (section 5.1.b) over a period of $t = 1000$ adimensional time and a 128×128 spatial grid. At the initial time (left figure in the first row) the relative vorticity is initialized with a monochromatic wave (solution of the deterministic BQG model). This monochromatic wave is still conserved by the stochastic system but a secondary circulation is triggered by the noise inhomogeneity (see Fig.3).
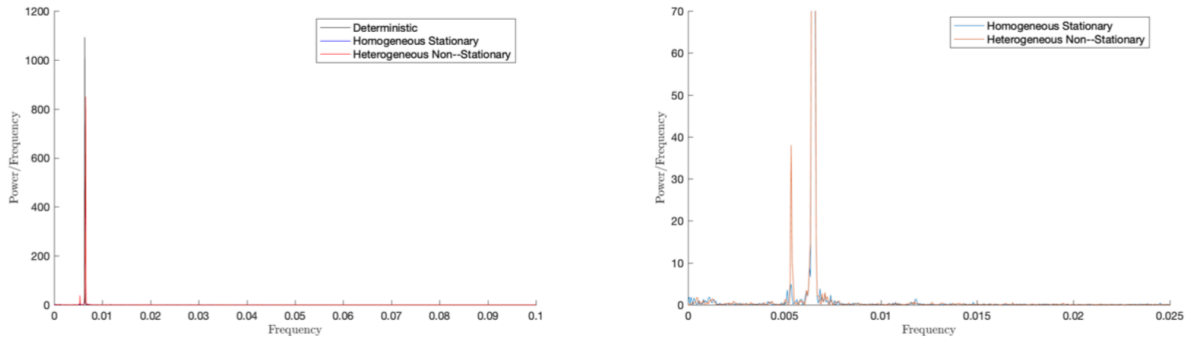
Figure 3: Fourier spectra of the temporal series of a fixed grid point for the deterministic BQG (black curve), BQG_{LU} with isotropic noise (blue curve) and BQG_{LU} with inhomogeneous noise (red curve); the right plot corresponds to a zoom of the left plots. The frequency of the initial monochromatic wave is conserved in all models (all the curves superimpose for this frequency). For the heterogeneous noise a second frequency associated to the secondary circulation emerges.

6   Conclusion

In this study, we have shown that a physically consistent stochastic representation, derived from conservation laws, introduces large-scale flow structuration caused by the action of the small-scale flow component. This effect is generated by a velocity transport component, associated to the inhomogeneity of the small-scale velocity fluctuations, that can be interpreted as a generalized Stokes drift – referred to in this work as Itô-Stokes drift. With this analogy, and proper assumptions we have shown that the derived stochastic dynamics encompasses the classical Craik-Leibovich system including a vortex force. This stochastic system, however, generalizes the dynamics capable of accounting for more complex interactions between the resolved current and the small-scale velocity fluctuations. The action of the velocity fluctuations towards large-scale flow structuration is demonstrated on a simple barotropic quasi-geostrophic model. The randomized system still conserves the Rossby wave structure, while introducing secondary vortices. On the contrary, a homogeneous isotropic noise conserves only the primary wave structure. This is a strong indication of the predominant role played by inhomogeneity of the small-scale velocity on shaping coherent large-scale structures in turbulent flows as first wonderfully intuited by Phillips (1977).

For simplicity, the Einstein summation convention is adopted in the following.

A   Quadratic (co-)variation

In stochastic calculus, the quadratic covariation (or cross-variance) of two processes X and Y is defined as:

$$\langle X, Y \rangle_t = \lim_{n \to 0} \sum_{i=1}^{p_n} (X_i^n - X_{i-1}^n)(Y_i^n - Y_{i-1}^n), \tag{38}$$

where $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ is a partition of the interval $[0, t]$ and this limit, if it exists, is defined in the sense of convergence in probability.

Assuming that X and Y are two continuous semimartingales, defined as $X_t = X_0 + A_t + M_t, Y_t = Y_0 + B_t + N_t$ with M, N martingales and A, B finite variation processes, then their quadratic covariation (38) exists, and is given by

$$\langle X, Y \rangle_t = \langle M, N \rangle_t. \tag{39}$$

In particular, the quadratic variation of a standard Brownian motion B (as a martingale) is given by $\langle B \rangle_t = t$.

The quadratic (co-)variations play an important role in the Itô calculus and its generalization of the chain rule. In particular, they are involved in the Itô integration by parts formula:

$$d_t(XY) = X d_t Y + Y d_t X + d\langle X, Y \rangle_t. \tag{40}$$

They are also involved in the Itô isometry, used to express the covariance of two Itô integrals:

$$\mathbb{E}\Big[\Big(\int_0^t f dM_s\Big)\Big(\int_0^t g dN_s\Big)\Big] = \mathbb{E}\Big[\int_0^t fg d\langle M, N \rangle_s\Big], \tag{41}$$

where f and g are two predictable processes such that $\int_0^t f^2 d\langle M, M \rangle_s$ and $\int_0^t g^2 d\langle N, N \rangle_s$ are integrable.

13

B Derivation of the stochastic transport operator

In this section, we show briefly how the stochastic transport operator $\mathbb{D}_t$ comes out. From basic knowledges of stochastic calculus, if θ is a deterministic function of class C^2 in space and of class C^1 in time, the differential of θ is given by the Itô formula (Kunita 1990). However, in our case the random scalar θ is a semimartingale (assumed to be continuous). This requires, therefore, to compute the differentiation of the composition of two stochastic processes. This is provided by the following generalized Itô-Wentzell formula (Kunita 1990) :

Theorem B.1 (Generalized Itô's formula I) *Let $\theta(\boldsymbol{x}, t), \boldsymbol{x} \in \Omega$ be a continuous C^2-process and a continuous C^1-semimartingale, let $\boldsymbol{X}_t$ be a continuous semimartingale with values in Ω. Then, $\theta(\boldsymbol{X}_t, t)$ is a continuous semimartingale satisfying*

$$d\theta(\boldsymbol{X}_t, t) = d_t\theta(\boldsymbol{X}_t, t) + \frac{\partial\theta}{\partial x_i}(\boldsymbol{X}_t, t)dX_t^i + \frac{1}{2}\frac{\partial^2\theta}{\partial x_i \partial x_j}(\boldsymbol{X}_t, t)d\Big\langle X^i, X^j\Big\rangle_t + d\Big\langle \frac{\partial\theta}{\partial x_i}(\boldsymbol{X}, \cdot), X^i\Big\rangle_t, \qquad (42)$$

where $d_t\theta(\boldsymbol{x}, t) \triangleq \theta(\boldsymbol{x}, t + dt) - \theta(\boldsymbol{x}, t)$ stands for the time increment.

For our stochastic flow (4), the first quadratic variation can be immediately determined from the definition (5), as follows

$$
\begin{aligned}
\Big\langle X^i, X^j\Big\rangle_t &= \Big\langle \int_0^\cdot \big(\sigma(\boldsymbol{X}_s, s)dB_s\big)^i, \int_0^\cdot \big(\sigma(\boldsymbol{X}_s, s)dB_s\big)^j\Big\rangle_t \\
&= \Big\langle \int_0^\cdot \int_\Omega \breve{\sigma}_{ik}(\boldsymbol{X}_s, \boldsymbol{y}, s)dB_s^k(\boldsymbol{y})d\boldsymbol{y}, \int_0^\cdot \int_\Omega \breve{\sigma}_{jl}(\boldsymbol{X}_s, \boldsymbol{z}, s)dB_s^l(\boldsymbol{z})d\boldsymbol{z}\Big\rangle_t \\
&= \int_0^t \int_{\Omega\times\Omega} \breve{\sigma}_{ik}(\boldsymbol{X}_s, \boldsymbol{y}, s)\breve{\sigma}_{jl}(\boldsymbol{X}_s, \boldsymbol{z}, s)\underbrace{d\Big\langle B^k(\boldsymbol{y}), B^l(\boldsymbol{z})\Big\rangle_s}_{\delta_{kl}\delta(\boldsymbol{y}-\boldsymbol{z})ds}d\boldsymbol{y}d\boldsymbol{z} \\
&= \int_0^t \underbrace{\int_\Omega \breve{\sigma}_{ik}(\boldsymbol{X}_s, \boldsymbol{y}, s)\breve{\sigma}_{kj}(\boldsymbol{X}_s, \boldsymbol{y}, s)d\boldsymbol{y}}_{\sigma_{ik}\sigma_{kj}(\boldsymbol{X}_s, s)}ds, \ \forall\, i, j = 1, \ldots, d,
\end{aligned}
\qquad (43)
$$

where δ denotes the Kronecker delta. This allows us to define

$$a_{ij}(\boldsymbol{X}_t, t) \triangleq \sigma_{ik}\sigma_{kj}(\boldsymbol{X}_t, t) = \frac{d\langle X^i, X^j\rangle_t}{dt}, \ \forall\, i, j = 1, \ldots, d. \qquad (44)$$

The last bracket in (42) is an additional term compared to the classical Itô formula. It describes the interaction between the stochastic flow and the gradient of the transported quantity. Writing the semimartingale θ in the form

$$d_t\theta(\boldsymbol{X}_t, t) = f(\boldsymbol{X}_t, t)dt + \int_\Omega g_k(\boldsymbol{X}_t, \boldsymbol{y}, t)dB_t^k(\boldsymbol{y})d\boldsymbol{y}, \qquad (45)$$

this quadratic variation can be specified in a similar way as in (43), namely

$$d\Big\langle \frac{\partial\theta}{\partial x_i}, X^i\Big\rangle_t = dt \int_\Omega \breve{\sigma}_{ij}(\boldsymbol{X}_t, \boldsymbol{y}, t)\frac{\partial g_j}{\partial x_i}(\boldsymbol{X}_t, \boldsymbol{y}, t)d\boldsymbol{y}. \qquad (46)$$

Substituting expressions (43) and (46) in Theorem B.1, we obtain the total variation of the scalar θ due to the stochastic flow (4). This reads

$$d\theta = d_t\theta + u_i\frac{\partial\theta}{\partial x_i}dt + \int_\Omega (\breve{\sigma}_{ij}dB_t^j)\frac{\partial\theta}{\partial x_i}d\boldsymbol{y} + \frac{1}{2}a_{ij}\frac{\partial^2\theta}{\partial x_i \partial x_j}dt + \int_\Omega \breve{\sigma}_{ij}\frac{\partial g_j}{\partial x_i}d\boldsymbol{y}dt. \qquad (47)$$

In particular, for a conserved scalar driven by the stochastic flow such that $d\theta(\boldsymbol{X}_t, t) = 0$, one may specify explicitly the expressions of f and g by identifying equation (47) and (45), namely

$$f = -u_i\frac{\partial\theta}{\partial x_i} - \frac{1}{2}a_{ij}\frac{\partial^2\theta}{\partial x_i \partial x_j} - \int_\Omega \breve{\sigma}_{ij}\frac{\partial g_j}{\partial x_i}d\boldsymbol{y}, \quad g_j = -\breve{\sigma}_{jk}\frac{\partial\theta}{\partial x_k}, \qquad (48)$$

since a semimartingale admits a unique decomposition (into a finite variation process and a martingale) (Kunita 1990). Substituting the expression of g in equation (47) for the last term, we have

$$
\begin{aligned}
\int_\Omega \breve{\sigma}_{ij}\frac{\partial g_j}{\partial x_i}d\boldsymbol{y} &= -\int_\Omega \breve{\sigma}_{ij}\breve{\sigma}_{jk}d\boldsymbol{y}\frac{\partial^2\theta}{\partial x_i \partial x_k} - \int_\Omega \breve{\sigma}_{ij}\frac{\partial\breve{\sigma}_{jk}}{\partial x_i}d\boldsymbol{y}\frac{\partial\theta}{\partial x_k} \\
&= -a_{ik}\frac{\partial^2\theta}{\partial x_i \partial x_k} - \underbrace{\int_\Omega \frac{\partial}{\partial x_i}(\breve{\sigma}_{ij}\breve{\sigma}_{jk})d\boldsymbol{y}}_{\frac{\partial a_{ik}}{\partial x_i}}\frac{\partial\theta}{\partial x_k} + \underbrace{\int_\Omega \breve{\sigma}_{jk}\frac{\partial\breve{\sigma}_{ij}}{\partial x_i}d\boldsymbol{y}}_{\sigma_{jk}\frac{\sigma_{ij}}{\partial x_i}}\frac{\partial\theta}{\partial x_k} \\
&= -\frac{1}{2}a_{ik}\frac{\partial^2\theta}{\partial x_i \partial x_k} - \frac{1}{2}\frac{\partial}{\partial x_i}\Big(a_{ik}\frac{\partial\theta}{\partial x_k}\Big) - \frac{1}{2}\frac{\partial a_{ik}}{\partial x_i}\frac{\partial\theta}{\partial x_k} + \sigma_{jk}\frac{\sigma_{ij}}{\partial x_i}\frac{\partial\theta}{\partial x_k}
\end{aligned}
\qquad (49)
$$

Sequentially substituting (49) in (47), the transport equation of θ reduces to

$$0 = \mathrm{d}_t\theta + \Big(u_k - \frac{1}{2}\frac{\partial a_{ik}}{\partial x_i} + \sigma_{jk}\frac{\sigma_{ij}}{\partial x_i}\Big)\frac{\partial\theta}{\partial x_k}\mathrm{d}t + (\sigma\mathrm{d}B_t)^k\frac{\partial\theta}{\partial x_k} - \frac{1}{2}\frac{\partial}{\partial x_i}\Big(a_{ik}\frac{\partial\theta}{\partial x_k}\Big)\mathrm{d}t$$

or, equivalently, in a vector form to

$$\mathbb{D}_t\theta \triangleq \mathrm{d}_t\theta + \Big(\big(\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} + \boldsymbol{\sigma}^T(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\big)\cdot\boldsymbol{\nabla}\theta - \frac{1}{2}\boldsymbol{\nabla}\cdot\big(\boldsymbol{a}\boldsymbol{\nabla}\theta\big)\Big)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\theta = 0. \tag{50}$$

The stochastic transport operator with incompressible constraint on the unresolved random component ensues immediately from this expression:

$$\mathbb{D}_t\theta \triangleq \mathrm{d}_t\theta + \Big(\big(\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}\big)\cdot\boldsymbol{\nabla}\theta - \frac{1}{2}\boldsymbol{\nabla}\cdot\big(\boldsymbol{a}\boldsymbol{\nabla}\theta\big)\Big)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\theta = 0. \tag{51}$$

C   Conversions between Itô and Stratonovich integrals

In this section, we give an equivalent expression of the stochastic transport operator (50) in the Stratonovich form. To this end, a general Stratonovich–Itô–integral conversion rule (Kunita 1990) is first adopted:

Theorem C.1 *If X and Y are two continuous semimartingales, the following Stratonovich integral is well defined:*

$$X_t \circ \mathrm{d}Y_t = X_t\mathrm{d}Y_t + \frac{1}{2}\mathrm{d}\langle X, Y\rangle_t. \tag{52}$$

Applying this formula component-wise to our stochastic flow (4) yields

$$\big(\sigma(\boldsymbol{X}_t, t)\circ\mathrm{d}B_t\big)^i = \int_\Omega \breve{\sigma}_{ij}(\boldsymbol{X}_t, \boldsymbol{y}, t)\mathrm{d}B_t^j(\boldsymbol{y})\mathrm{d}\boldsymbol{y} + \underbrace{\frac{1}{2}\int_\Omega\big\langle\mathrm{d}\breve{\sigma}_{ij}(\boldsymbol{X}_t, \boldsymbol{y}, t), \mathrm{d}B_t^j(\boldsymbol{y})\big\rangle\mathrm{d}\boldsymbol{y}}_{I}. \tag{53}$$

Since $\boldsymbol{\sigma}$ is assumed to be deterministic in (5), the first term in I can be determined, by the classical Itô formula (i.e. (42) without the last quadratic variation process), as follows:

$$\mathrm{d}\breve{\sigma}_{ij} = \frac{\partial\breve{\sigma}_{ij}}{\partial t}\mathrm{d}t + \frac{\partial\breve{\sigma}_{ij}}{\partial x_k}\underbrace{\Big(u^k\mathrm{d}t + \int_\Omega\breve{\sigma}_{kl}(\boldsymbol{X}_t, \boldsymbol{z}, t)\mathrm{d}B_t^l(\boldsymbol{z})\mathrm{d}\boldsymbol{z}\Big)}_{\mathrm{d}X_t^k} + \frac{1}{2}\frac{\partial^2\breve{\sigma}_{ij}}{\partial x_k\partial x_l}\mathrm{d}\langle X^k, X^l\rangle_t. \tag{54}$$

Substituting (54) for the term I in (53), we get

$$I = \frac{1}{2}\int_{\Omega\times\Omega}\frac{\partial\breve{\sigma}_{ij}}{\partial x_k}\breve{\sigma}_{kl}\underbrace{\mathrm{d}\langle B^l(\boldsymbol{z}), B^j(\boldsymbol{y})\rangle_t}_{\delta_{jl}\delta(\boldsymbol{y}-\boldsymbol{z})\mathrm{d}t}\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{z} = \frac{1}{2}\frac{\partial\sigma_{ij}}{\partial x_k}\sigma_{kj}\mathrm{d}t = \frac{1}{2}\Big(\frac{\partial a_{ik}}{\partial x_k} - \sigma_{ij}\frac{\partial\sigma_{kj}}{\partial x_k}\Big)\mathrm{d}t. \tag{55}$$

Subsequently, substituting (55) in (53), then in (4), we deduce an equivalent Stratonovich representation of our stochastic flow that reads in vector form as

$$\mathrm{d}\boldsymbol{X}_t = \Big(\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} + \frac{1}{2}\boldsymbol{\sigma}^T(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\Big)\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t. \tag{56}$$

Afterwards, for a random tracer θ (which is now assumed to be a smoother field than in Appendix B – more precisely it is assumed to be C^3 in space and C^2 in time) advected by a Stranotovich flow, one can apply the following generalized Itô's formula (Kunita 1990):

Theorem C.2 (Generalized Itô's formula II) *Let $\theta(\boldsymbol{x}, t), \boldsymbol{x}\in\Omega$ be a continuous C^3-process and a continuous C^2-semimartingale, let $\boldsymbol{X}_t$ be a continuous semimartingale with values in Ω. Then, the following formula is satisfied:*

$$\mathrm{d}\theta(\boldsymbol{X}_t, t) = \mathrm{d}_t\circ\theta(\boldsymbol{X}_t, t) + \frac{\partial\theta}{\partial x_i}(\boldsymbol{X}_t, t)\circ\mathrm{d}X_t^i, \tag{57}$$

where $\mathrm{d}_t\circ\theta(\boldsymbol{x}, t) \triangleq \theta(\boldsymbol{x}, t+\frac{\mathrm{d}t}{2}) - \theta(\boldsymbol{x}, t-\frac{\mathrm{d}t}{2})$ stands for the centered time increment.

Substituting (56) in (57), we deduce a Stratonovich representation of the stochastic transport operator (50), that is

$$\mathbb{D}_t^S\theta \triangleq \mathrm{d}_t\circ\theta + \Big(\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} + \frac{1}{2}\boldsymbol{\sigma}^T(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\Big)\cdot\boldsymbol{\nabla}\theta\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\theta = 0. \tag{58}$$

For incompressible flows, the Stratonovich transport operator reads

$$\mathbb{D}_t^S\theta \triangleq \mathrm{d}_t\circ\theta + \Big(\boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}\Big)\cdot\boldsymbol{\nabla}\theta\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\theta = 0. \tag{59}$$

On the other hand, if we impose a Stratonovich decomposition at the beginning as given in (29), by applying Theorem C.2 directly, the transport equation can be simply written as

$$\mathrm{d}_t\circ\theta + (\boldsymbol{u}'\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\theta = 0. \tag{60}$$

15

Nevertheless, one must keep in mind that numerous additional terms are hidden behind the Stratonovich integral (i.e. in the two brackets of Theorem (B.1)). For instance, to interpret the smooth flow component $\boldsymbol{u}'$, one may first derive the corresponding Itô form of the Stratonovich flow (29), that is

$$\mathrm{d}\boldsymbol{X}_t = \left(\boldsymbol{u}' + \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} - \frac{1}{2}\boldsymbol{\sigma}^T(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\right)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \tag{61}$$

which can be shown in a similar way as in (52)-(56). Then, identifying this expression with (4), we deduce

$$\boldsymbol{u}' = \boldsymbol{u} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} + \frac{1}{2}\boldsymbol{\sigma}^T(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}). \tag{62}$$

Therefore, for incompressible noise, we have $\boldsymbol{u}' = \boldsymbol{u}^*$.

D Derivation of the BQG$_{\mathbf{LU}}$

In this section, the derivation of governing equations for stochastic QG flows is performed with a similar strategy as in the classical framework. In particular, the physical arguments used in the classical derivation are also adopted. We start from the 2D (rotating) shallow water system under location uncertainty given in (Mémin 2014), that is

$$\mathbb{D}_t\boldsymbol{u} + f\boldsymbol{k} \times (\boldsymbol{u}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = -\boldsymbol{\nabla}(\phi'\mathrm{d}t + \mathrm{d}\Phi_t), \tag{63a}$$

$$\mathbb{D}_t\phi' + (c_0^2 + \phi')\boldsymbol{\nabla}\cdot\boldsymbol{u}\mathrm{d}t = 0, \tag{63b}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0, \tag{63c}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot\boldsymbol{a} = 0, \tag{63d}$$

where $\boldsymbol{k} = (0,0,1)^T$ and $\phi' = gh'$ stands for a perturbation of the geopotential $\phi = c_0^2 + \phi', c_0^2 = gH$, associated to a small perturbation of the surface height h' w.r.t. a background height H. The quantity $\mathrm{d}\Phi_t$ is a rescaled random pressure contribution. We consider the following scalings

$$\boldsymbol{x} = L\boldsymbol{x}_1, \ \boldsymbol{u} = U\boldsymbol{u}_1, \ t = (L/U)t_1, \tag{64a}$$

$$f = f_0 f_1, \ f_1 = 1 + \frac{\beta y}{f_0} = 1 + Ro\beta_1 y_1, \tag{64b}$$

$$\phi' = f_0 U L \phi_1, \tag{64c}$$

$$\boldsymbol{a} = \epsilon U L \boldsymbol{a}_1, \tag{64d}$$

$$\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \sqrt{\epsilon}L(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1, \tag{64e}$$

$$\mathrm{d}\Phi_t = \sqrt{\epsilon}f_0 L^2(\mathrm{d}\Phi_t)_1, \tag{64f}$$

where the capital numbers stand for the typical values of physical quantities, the numbers with subscript 1 are denoted as dimensionless quantities, $Ro = U/(f_0 L)$ is the Rossby number and ϵ is a special scaling for the amplitude of the variance tensor. In this work, we set it to be $\mathscr{O}(1)$ [or $\epsilon Ro = \mathscr{O}(Ro)$]. Substituting these expressions in (63), we obtain the following non-dimensional system

$$Ro\mathbb{D}_{t_1}^\epsilon \boldsymbol{u}_1 + f_1\boldsymbol{k} \times \left(\boldsymbol{u}_1\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1\right) = -\boldsymbol{\nabla}_1\left(\phi_1\mathrm{d}t_1 + \sqrt{\epsilon}(\mathrm{d}\Phi_t)_1\right), \tag{65a}$$

$$Ro\mu_0^2\mathbb{D}_{t_1}^\epsilon \phi_1 + \left(1 + Ro\mu_0^2\phi_1\right)\boldsymbol{\nabla}_1\cdot\boldsymbol{u}_1\mathrm{d}t_1 = 0, \tag{65b}$$

$$\boldsymbol{\nabla}_1\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1 = 0, \tag{65c}$$

$$\boldsymbol{\nabla}_1\cdot\boldsymbol{\nabla}_1\cdot\boldsymbol{a}_1 = 0, \tag{65d}$$

where $\mathbb{D}_{t_1}^\epsilon\theta \triangleq \mathrm{d}_{t_1}\theta + (\boldsymbol{u}_1^\star\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1)\cdot\boldsymbol{\nabla}_1\theta - \frac{\epsilon}{2}\boldsymbol{\nabla}_1\cdot(\boldsymbol{a}_1\boldsymbol{\nabla}_1\theta)\mathrm{d}t_1, \boldsymbol{u}_1^\star = \boldsymbol{u}_1 - \frac{\epsilon}{2}\boldsymbol{\nabla}_1\cdot\boldsymbol{a}_1$ and $\mu_0 \triangleq f_0 L/c_0$ defined as a scaling constant. We then expand $\boldsymbol{u}_1$ and ϕ_1 in power series of (small) Rossby number

$$\boldsymbol{u}_1 = \boldsymbol{u}_1^{(0)} + Ro\boldsymbol{u}_1^{(1)} + \cdots, \quad \phi_1 = \phi_1^{(0)} + Ro\phi_1^{(1)} + \cdots. \tag{66}$$

Substituting these expressions in (65) while using (64b) and keeping only the zeroth and first orders quantities in the Rossby number,

$$Ro\mathbb{D}_{t_1}^{\epsilon(0)}\boldsymbol{u}_1^{(0)} + \boldsymbol{k} \times \left(\boldsymbol{u}_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1\right) + Ro\boldsymbol{k} \times \left(\beta_1 y_1\left(\boldsymbol{u}_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1\right) + \boldsymbol{u}_1^{(1)}\mathrm{d}t_1\right)$$
$$= -\boldsymbol{\nabla}_1\left(\phi_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\mathrm{d}\Phi_t)_1\right) - Ro\boldsymbol{\nabla}_1\phi_1^{(1)}\mathrm{d}t_1 + \mathscr{O}(Ro^2), \tag{67a}$$

$$Ro\mu_0^2\mathbb{D}_{t_1}^{\epsilon(0)}\phi_1^{(0)} + Ro\boldsymbol{\nabla}_1\cdot\boldsymbol{u}_1^{(1)}\mathrm{d}t_1 + \left(1 + Ro\mu_0^2\phi_1^{(0)}\right)\boldsymbol{\nabla}_1\cdot\boldsymbol{u}_1^{(0)}\mathrm{d}t_1 = \mathscr{O}(Ro^2), \tag{67b}$$

where $\mathbb{D}_{t_1}^{\epsilon(0)}\theta \triangleq \mathrm{d}_{t_1}\theta + \left((\boldsymbol{u}_1^{(0)} - \frac{\epsilon}{2}\boldsymbol{\nabla}_1\cdot\boldsymbol{a}_1)\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1\right)\cdot\boldsymbol{\nabla}_1\theta - \frac{\epsilon}{2}\boldsymbol{\nabla}_1\cdot(\boldsymbol{a}_1\boldsymbol{\nabla}_1\theta)\mathrm{d}t_1$. The zeroth order equations reduce to

$$\boldsymbol{u}_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1 = \boldsymbol{\nabla}_1^\perp\left(\phi_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\mathrm{d}\Phi_t)_1\right), \tag{68a}$$

$$\boldsymbol{\nabla}_1 \cdot \boldsymbol{u}_1^{(0)} = 0. \tag{68b}$$

From (65c), equation (68a) reduces to

$$\boldsymbol{u}_1^{(0)} = \boldsymbol{\nabla}_1^{\perp}\phi_1^{(0)}, \quad (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1 = \boldsymbol{\nabla}_1^{\perp}(\mathrm{d}\Phi_t)_1. \tag{69}$$

The first order equations reduce to

$$\mathbb{D}_{t_1}^{\epsilon(0)}\boldsymbol{u}_1^{(0)} + \boldsymbol{k} \times \left(\beta_1 y_1 \big(\boldsymbol{u}_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1\big) + \boldsymbol{u}_1^{(1)}\mathrm{d}t_1\right) = -\boldsymbol{\nabla}_1\phi_1^{(1)}\mathrm{d}t_1, \tag{70a}$$

$$\mu_0^2\mathbb{D}_{t_1}^{\epsilon(0)}\phi_1^{(0)} + \boldsymbol{\nabla}_1 \cdot \boldsymbol{u}_1^{(1)}\mathrm{d}t_1 = 0. \tag{70b}$$

Taking the curl $(\boldsymbol{k} \cdot \boldsymbol{\nabla}_1 \times)$ of equation (70a), we deduce an evolution law for the relative vorticity, defined as $\zeta_1^{(0)} \triangleq \boldsymbol{k} \cdot \boldsymbol{\nabla}_1 \times \boldsymbol{u}_1^{(0)} = \frac{\partial v_1^{(0)}}{\partial x_1} - \frac{\partial u_1^{(0)}}{\partial y_1}$. It reads

$$\mathbb{D}_{t_1}^{\epsilon(0)}\zeta_1^{(0)} + \beta_1\left(v_1^{(0)}\mathrm{d}t_1 + \sqrt{\epsilon}(\sigma\mathrm{d}B_t)_1^y\right) + \boldsymbol{\nabla}_1 \cdot \boldsymbol{u}_1^{(1)}\mathrm{d}t_1 = \frac{\epsilon}{2}\gamma\Big(\boldsymbol{a}_1, \boldsymbol{u}_1^{(0)}\Big)\mathrm{d}t_1 + \sqrt{\epsilon}\eta\Big((\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1, \boldsymbol{u}_1^{(0)}\Big), \tag{71a}$$

$$\gamma\Big(\boldsymbol{a}_1, \boldsymbol{u}_1^{(0)}\Big) = \mathrm{tr}\Big(\boldsymbol{\nabla}_1^{\perp}\big(\boldsymbol{\nabla}_1 \cdot \boldsymbol{a}_1\big)^T\boldsymbol{\nabla}_1\big(\boldsymbol{u}_1^{(0)}\big)^T\Big) + \boldsymbol{\nabla}_1 \cdot \Big(\frac{\partial \boldsymbol{a}_1}{\partial x_1}\boldsymbol{\nabla}_1 v_1^{(0)} - \frac{\partial \boldsymbol{a}_1}{\partial y_1}\boldsymbol{\nabla}_1 u_1^{(0)}\Big), \tag{71b}$$

$$\eta\Big((\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1, \boldsymbol{u}_1^{(0)}\Big) = \mathrm{tr}\Big(\boldsymbol{\nabla}_1^{\perp}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_1^T\boldsymbol{\nabla}_1(\boldsymbol{u}_1^{(0)})^T\Big). \tag{71c}$$

Note that the following vector calculus identity has been used here several times :

$$\boldsymbol{k} \cdot \boldsymbol{\nabla} \times \big((\boldsymbol{A} \cdot \boldsymbol{\nabla})\boldsymbol{B}\big) = \boldsymbol{\nabla}^{\perp} \cdot \big((\boldsymbol{A} \cdot \boldsymbol{\nabla})\boldsymbol{B}\big) = \mathrm{tr}\big(\boldsymbol{\nabla}^{\perp}\boldsymbol{A}^T\boldsymbol{\nabla}\boldsymbol{B}^T\big) + (\boldsymbol{A} \cdot \boldsymbol{\nabla})(\boldsymbol{\nabla}^{\perp} \cdot \boldsymbol{B}), \ \forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^2.$$

We then go back to the dimensional equation. The zeroth order quantities satisfy the scalings (64), while (66), we deduce the scaling of the first order velocity such as $\boldsymbol{u}^{(1)} = Ro U \boldsymbol{u}_1^{(1)}$. Thus, the dimensional version of equation (71a) and equation (70b) can be written, respectively, as

$$\mathbb{D}_t^{(0)}\zeta^{(0)} + \beta\Big(v^{(0)}\mathrm{d}t + (\sigma\mathrm{d}B_t)^y\Big) + f_0\boldsymbol{\nabla} \cdot \boldsymbol{u}^{(1)}\mathrm{d}t = \frac{1}{2}\gamma\Big(\boldsymbol{a}, \boldsymbol{u}^{(0)}\Big)\mathrm{d}t + \eta\Big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \boldsymbol{u}^{(0)}\Big), \tag{72a}$$

$$\mathbb{D}_t^{(0)}\phi^{(0)} + c_0^2\boldsymbol{\nabla} \cdot \boldsymbol{u}^{(1)}\mathrm{d}t = 0, \tag{72b}$$

where $\mathbb{D}_t^{(0)}\theta \triangleq \mathrm{d}_t\theta + \big((\boldsymbol{u}^{(0)} - \frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{a})\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big) \cdot \boldsymbol{\nabla}\theta - \frac{1}{2}\boldsymbol{\nabla} \cdot (\boldsymbol{a}\boldsymbol{\nabla}\theta)\mathrm{d}t$. Differentiating between equation (72a) and equation (72b) multiplied by f_0/c_0^2, as well as taking account of the Coriolis evolution, we obtain a single equation

$$\mathbb{D}_t^{(0)}\Big(\zeta^{(0)} - \frac{f_0}{c_0^2}\phi^{(0)} + f_0 + \beta y\Big) = \frac{1}{2}\gamma\Big(\boldsymbol{a}, \boldsymbol{u}^{(0)}\Big)\mathrm{d}t + \eta\Big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \boldsymbol{u}^{(0)}\Big) - \beta\boldsymbol{\nabla} \cdot \boldsymbol{a}_{.y}\mathrm{d}t. \tag{73}$$

According to the incompressibility of $\boldsymbol{u}^{(0)}$ and $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ from (69), we introduce, respectively, a resolved streamfunction $\psi^{(0)} = \phi^{(0)}/f_0$ and an unresolved one $\varphi\mathrm{d}B_t = \mathrm{d}\Phi_t/f_0$ such that $\boldsymbol{u}^{(0)} = \boldsymbol{\nabla}^{\perp}\psi^{(0)}$ and $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^{\perp}\varphi\mathrm{d}B_t$. Thus, the previous equation (73) may be re-written as

$$\mathbb{D}_t^{(0)}\Big(\nabla^2\psi^{(0)} - \frac{f_0^2}{c_0^2}\psi^{(0)} + f\Big) = \mathrm{d}R_t^{(0)}, \tag{74}$$

where $\mathrm{d}R_t^{(0)} \triangleq \frac{1}{2}\gamma\Big(\boldsymbol{a}, \boldsymbol{u}^{(0)}\Big)\mathrm{d}t + \eta\Big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \boldsymbol{u}^{(0)}\Big) - (\boldsymbol{\nabla} \cdot \boldsymbol{a}) \cdot \boldsymbol{\nabla}f\mathrm{d}t$.

Finally, when the Rossby number tends to zero, we obtain a stochastic advection-diffusion equation with sources and sinks for the PV, that is

$$\mathbb{D}_t q = \mathrm{d}R_t, \quad q = \nabla^2\psi - \frac{1}{L_R^2}\psi + f, \tag{75}$$

where $L_R \triangleq \frac{\sqrt{gH}}{f_0} = \frac{c_0}{f_0}$.

On the other hand, if we use the Stratonovich transport operator (59) to derive the governing equations, the equivalent Stratonovich representation of the system (75) can be written as

$$\mathbb{D}_t^S q = \underbrace{\frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\nabla}^{\perp}(\boldsymbol{\nabla} \cdot \boldsymbol{a})^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big)\mathrm{d}t - \mathrm{tr}\Big(\boldsymbol{\nabla}^{\perp}(\boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t)^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big) - \frac{1}{2}(\boldsymbol{\nabla} \cdot \boldsymbol{a}) \cdot \boldsymbol{\nabla}f\mathrm{d}t}_{\mathrm{d}R_t^S} \tag{76a}$$

$$q = \nabla^2\psi - \frac{1}{L_R^2}\psi + f, \tag{76b}$$

$$\boldsymbol{u} = \boldsymbol{\nabla}^{\perp}\psi, \quad \boldsymbol{\nabla} \cdot \boldsymbol{a} = \boldsymbol{\nabla}^{\perp}\tilde{\psi}, \quad \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^{\perp}\varphi\mathrm{d}B_t. \tag{76c}$$

E   Energetics budget of the BQG$_{\mathbf{LU}}$

In this section, we show the conservation of total energy for the BQG$_{\text{LU}}$, with a similar strategy as in the classical deterministic framework (Pedlosky 1992). For ease of understanding, we use the previous Stratonovich notation (76) in the following.

Recall that the total energy in a BQG system is defined as (Pedlosky 1992)

$$E = \int_\Omega \frac{1}{2}\|\boldsymbol{\nabla}\psi\|^2 + \frac{1}{2}F\psi^2, \ \ F \triangleq \frac{1}{L_R^2}. \tag{77}$$

We now compute its time evolution in the case of BQG-LU (76) :

$$
\begin{aligned}
\mathrm{d}_t E &= \int_\Omega (\mathrm{d}_t \boldsymbol{\nabla}\psi) \cdot \boldsymbol{\nabla}\psi + F\psi \mathrm{d}_t\psi \\
&= -\int_\Omega \psi \boldsymbol{\nabla}\cdot(\mathrm{d}_t\psi) + \int_\Omega F\psi \mathrm{d}_t\psi \\
&= -\int_\Omega \psi \mathrm{d}_t(\nabla^2\psi - F\psi) = -\int_\Omega \psi \mathrm{d}_t q, \ \text{via (76b)} \\
&= \int_\Omega \psi(\boldsymbol{u}^\star \mathrm{d}t + \boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}q - \int_\Omega \psi \mathrm{d}R_t^S, \ \text{via (76a)} \\
&= \int_\Omega \psi \boldsymbol{\nabla}\cdot\Big((\boldsymbol{u}^\star \mathrm{d}t + \boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)q\Big) - \int_\Omega \psi \mathrm{d}R_t^S, \ \text{via (76c)} \\
&= -\int_\Omega \boldsymbol{\nabla}\psi\cdot(\boldsymbol{u}^\star \mathrm{d}t + \boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)q - \int_\Omega \psi \mathrm{d}R_t^S \\
&= \underbrace{-\int_\Omega \boldsymbol{\nabla}\psi\cdot(-\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}\,\mathrm{d}t + \boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)q}_{I} - \int_\Omega \psi \mathrm{d}R_t^S, \ \text{via (76c)},
\end{aligned}
\tag{78}
\tag{79}
$$

where equations (78) and (79) are obtained by integration by parts within a close impermeable boundary condition (i.e. uniform ψ along the boundary) or with doubly periodic boundaries. Notice that until this step, we show that without any noise, the classical QG system conserves well the total energy.

Let us expand the contribution of the random source/sink term as follows :

$$
\begin{aligned}
\int_\Omega \psi \mathrm{tr}\Big(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)^T \boldsymbol{\nabla}\boldsymbol{u}^T\Big) &= \int_\Omega \psi \partial_i^\perp(\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}u_i \\
&= \int_\Omega \boldsymbol{\nabla}\cdot\Big(\psi\partial_i^\perp(\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)u_i\Big), \ \text{via (76c)} \\
&= -\int_\Omega \boldsymbol{\nabla}\psi\cdot\partial_i^\perp(\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t)u_i \\
&= -\int_\Omega (\boldsymbol{\nabla}\psi)^T(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp)\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t \\
&= \int_\Omega \boldsymbol{\nabla}^\perp\cdot\Big(\boldsymbol{u}(\boldsymbol{\nabla}\psi)^T\Big)\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t \\
&= \int_\Omega \boldsymbol{\nabla}\psi\cdot\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t\nabla^2\psi + \underbrace{\int_\Omega\Big(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp(\boldsymbol{\nabla}\psi)^T\Big)\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t}_{II},
\end{aligned}
\tag{80}
\tag{81}
$$

where $\partial_1^\perp v \triangleq -\partial_y$, $\partial_2^\perp \triangleq \partial_x$ and equations (80) and (81) are obtained by integration by parts. Moreover, the last term in the previous equation is null through (76c). Indeed,

$$
\begin{aligned}
II &= \int_\Omega\Big(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp(\boldsymbol{\nabla}\psi)^T\Big)\boldsymbol{\nabla}^\perp\varphi\circ \mathrm{d}B_t \\
&= -\int_\Omega \boldsymbol{\nabla}^\perp\cdot\Big(\boldsymbol{u}\cdot\boldsymbol{\nabla}^\perp(\boldsymbol{\nabla}\psi)^T\Big)\varphi\circ \mathrm{d}B_t \\
&= \int_\Omega \boldsymbol{\nabla}^\perp\cdot\Big(\boldsymbol{u}^\perp\cdot\boldsymbol{\nabla}(\boldsymbol{\nabla}\psi)^T\Big)\varphi\circ \mathrm{d}B_t \\
&= \int_\Omega \mathrm{tr}\Big(\boldsymbol{\nabla}^\perp(\boldsymbol{u}^\perp)^T\boldsymbol{\nabla}(\boldsymbol{u}^\perp)^T\Big) + (\boldsymbol{u}^\perp\cdot\boldsymbol{\nabla})(\boldsymbol{\nabla}^\perp\cdot\boldsymbol{\nabla}\psi) = 0
\end{aligned}
\tag{82}
$$

Similarly, one can show the result for the other source/sink term (of bounded variation) in (76a), that is

$$\int_\Omega \psi \mathrm{tr}\Big(\boldsymbol{\nabla}^\perp(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\boldsymbol{\nabla}\boldsymbol{u}^T\Big) = \int_\Omega \boldsymbol{\nabla}\psi\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})\nabla^2\psi. \tag{83}$$

This allows us to deduce the contribution of whole sources and sinks to the energetic budget as

$$-\int_\Omega \psi \mathrm{d}R_t^S = \int_\Omega \boldsymbol{\nabla}\psi\cdot\boldsymbol{\sigma}\circ \mathrm{d}\boldsymbol{B}_t\nabla^2\psi - \frac{\mathrm{d}t}{2}\int_\Omega \boldsymbol{\nabla}\psi\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})\nabla^2\psi + \frac{\mathrm{d}t}{2}\int_\Omega \psi(\boldsymbol{\nabla}\cdot\boldsymbol{a})\cdot\boldsymbol{\nabla}f. \tag{84}$$

On the other hand, we develop the contribution of the stochastic transport operator like

$$I = -\int_\Omega \boldsymbol{\nabla}\psi \cdot \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t (\nabla^2\psi - F\psi + f) + \frac{\mathrm{d}t}{2} \int_\Omega \boldsymbol{\nabla}\psi \cdot (\boldsymbol{\nabla}\cdot\boldsymbol{a})(\nabla^2\psi - F\psi) - \frac{\mathrm{d}t}{2}\int_\Omega \psi(\boldsymbol{\nabla}\cdot\boldsymbol{a})\cdot\boldsymbol{\nabla}f. \quad (85)$$

Thus, the evolution of the total energy reduces to

$$\begin{aligned}
\mathrm{d}_t E &= \int_\Omega F\psi\boldsymbol{\nabla}\psi \cdot \left(\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}\mathrm{d}t\right) - \int_\Omega f\boldsymbol{\nabla}\psi\cdot\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t \\
&= \int_\Omega \boldsymbol{\nabla}\cdot\left(F\psi\boldsymbol{\nabla}^\perp\psi\right)(\varphi\circ\mathrm{d}\boldsymbol{B}_t - \tilde{\psi}\mathrm{d}t) - \int_\Omega \boldsymbol{\nabla}\cdot(f\boldsymbol{\nabla}^\perp\psi)\varphi\circ\mathrm{d}\boldsymbol{B}_t \\
&= \int_\Omega \left(F\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}^\perp\psi + F\psi\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}^\perp\psi)\right)(\varphi\circ\mathrm{d}\boldsymbol{B}_t - \tilde{\psi}\mathrm{d}t) - \int_\Omega (\boldsymbol{\nabla}f\cdot\boldsymbol{\nabla}^\perp\psi)\varphi\circ\mathrm{d}\boldsymbol{B}_t \\
&= -\beta\int_\Omega \frac{\partial\psi}{\partial x}\varphi\circ\mathrm{d}\boldsymbol{B}_t.
\end{aligned}$$

Hence, the BQG-LU system conserves the total energy in a f–plane. The energy can be considered to be conserved also in a β–plane since $|\beta L| \ll |f_0|$.

References

Andrejczuk, M., F. Cooper, S. Juricke, T. Palmer, A. Weisheimer, and L. Zanna, 2016: Oceanic stochastic parameterization in a seasonal forecast system. *Month. Weath. R.*, **144 (1867-1875)**.

Andrews, D., and M. McIntyre, 1978: An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid. Mech.*, **89 (4)**, 609–646.

Arakawa, A., and V. Lamb, 1977: Computational design of the basic dynamical process of the UCLA general circulation mode. *Methods Computational Physics*, **17**, 173–265.

Ardhuin, F., N. Rascle, and K. Belibassakis, 2008: Explicit wave-averaged primitive equations using a generalized Lagrangian mean,. *Ocean Modell.*, **20 (1)**, 35–60.

Boussinesq, J., 1877: Mémoires présentés par divers savants à l'Académie des Sciences, 23 (1): 1–680.

Buizza, R., M. Miller, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal Royal Meteorological Society*, **125**, 2887–2908.

Chandramouli, P., D. Heitz, S. Laizet, and E. Mémin, 2018: Coarse large-eddy simulations in a transitional wake flow with flow models under location uncertainty. *Comp. & Fluids*, **168**, 170–189.

Chapron, B., P. Dérian, E. Mémin, and V. Resseguier, 2018: Large-scale flows under location uncertainty: a consistent stochastic framework. *QJRMS*, **144 (710)**, 251–260.

Craik, A., and S. Leibovich, 1976: Rational model for langmuir circulations. *J. Fluid Mech.*, **73**, 401–426.

Franzke, C., and A. Majda, 2005: Low-order stochastic mode reduction for a prototype atmospheric gcm. *J. Atmos. Sci.*, **63**, 457–479.

Glazunov, A., 2010: On the effect that the direction of geostrophic wind has on turbulence and quasiordered large-scale structures in the atmospheric boundary layer. *Izvestiya, Atmospheric and Oceanic Physics*, **46 (6)**, 727–747.

Gottlieb, S., 2005: On high order strong stability preserving runge-kutta and multi step time discretizations. *J. Sci. Comput.*, **25 (1)**, 105–128.

Gugole, F., and C. Franzke, 2019: Numerical development and evaluation of an energy conserving conceptual stochastic climate model. *Mathematics of climate and weather forecasting*, **5 (1)**, 45–64.

Haney, S., and W. Young, 2017: Radiation of internal waves from groups of surface gravity waves. *J. Fluid Mech.*, **829**, 280–303.

Harcourt, R., 2015: An improved second-moment closure model of langmuir turbulence. *J. Phys. Oceanogr.*, **45**, 84–103.

Harcourt, R., and E. D'Asaro, 2008: Large-eddy simulation of langmuir turbulence in pure wind seas. *J. Phys. Oceanogr., 38(7),*, **38 (7)**, 1542–1562.

Hasselmann, K., 1971: On the mass and momentum transfer between short gravity waves and larger-scale motions. *J. Fluid Mech.*, **50**, 189–201.

Hasselmann, K., 1976: Stochastic climate models part I theory. *Tellus*, **28**, 473–485.

Haworth, D., and S. Pope, 1986: A generalized Langevin model for turbulent flows. *Phys. Fluids*, **29**, 387–405.

Holm, D., 2015: Variational principles for stochastic fluid dynamics. *Proc. R. Soc. A*, **471 (20140963)**.

Holm, D., 2019: Stochastic closures for wave–current interaction dynamics. Tech. rep., arXiv:1905.01930.

Holm, D. D., 1996: The ideal craik-leibovich equations. *Physica-D*, **98**, 415–441.

Kadri Harouna, S., and E. Mémin, 2017: Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling. *Computers & Fluids*, **156**, 456–469, doi:10.1016/j.compfluid.2017.08.017, URL https://hal.inria.fr/hal-01394780.

Kunita, H., 1990: *Stochastic flows and stochastic differential equations*. Cambridge University Press.

Lane, E., J. Restrepo, and J. McWilliams, 2007: Wave-current interaction: A comparison of radiation-stress and vortex-force representations. *J. Phys. Oceanogr.*, **37 (5)**, 1122–1141.

Leibovich, S., 1980: On wave-current interaction theories of langmuir circulations. *J. Fluid Mech.*, **99 (4)**, 715–724.

Leith, C., 1990: Stochastic backscatter in a subgrid-scale model: plane shear mixing layer. *Phys. of Fluids*, **2 (3)**, 1521–1530.

Liu, J., J.-H. Liang, J. McWilliams, P. Sullivan, Y. Fan, and Q. Chen, 2018: Effect of planetary rotation on oceanic surface boundary layer turbulence. *Journ of phys. Oceanogr.*, **48**, 2057–2080.

Longuet-Higgins, M., and R. Stewart, 1964: Radiation stresses in water waves: A physical discussion, with applications. *Deep Sea Res. Oceanogr. Abstr.*, **11**, 529–562.

McWilliams, J., 1984: The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.*, **146 (1)**, 21.

McWilliams, J., J. Restrepo, and E. Lane, 2004: An asymptotic theory for the interaction of waves and currents in coastal waters. *J. of Fluid Mech.*, **511**, 135–178.

McWilliams, J., P. Sullivan, and C.-H. Moeng, 1997: Langmuir turbulence in the ocean. *J. Fluid Mech.*, **334**, 1–30.

Mellor, G., 2003: The three-dimensional current and surface wave equations. *J. Phys. Oceanogr.*, **33 (9)**, 1978–1989.

Mellor, G., 2016: On theories dealing with the interaction of surface waves and ocean circulation. *J. Geophys. Res.*, **121**, 4474–4486.

Mémin, E., 2014: Fluid flow dynamics under location uncertainty. *Geophys. & Astro. Fluid Dyn.*, **108 (2)**, 119–146, doi:10.1080/03091929.2013.836190.

Meneveau, C., and J. Katz, 2000: Scale-invariance and turbulence models for large-eddy simulation. *Annu. Rev. Fluid. Mech*, **32**, 1–32.

Pavliotis, G., and A. Stuart, 2008: *Multiscale methods: averaging and homogenization*. Science & Business Media, Springer.

Pedlosky, J., 1992: *Geophysical Fluid Dynamics*. Springer New York.

Phillips, O. M., 1977: *The Dynamics of the Upper Ocean*. Cambridge Univ. Press, Cambridge, U. K.

Pinier, B., E. Mémin, S. Laizet, and R. Lewandowski, 2019: Stochastic flow approach to model the mean velocity profile of wall-bounded flows. *Phys. Rev. E*, **99 (6)**, 063 101.

Pope, S., 1994: Lagrangian pdf methods for turbulent flows. *Annu. Rev. Fluid Mech.*

Pope, S., 2000: *Turbulent flows*. Cambridge University Press.

Porta Mana, P., and L. Zanna, 2014: Toward a stochastic parametrization of ocean mesoscale eddies. *Ocean Modelling*, **79 (1-20)**.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery, 2007: *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press.

Reeks, M., 1983: The transport of discrete particles in inhomogeneous turbulence. *J. Aerosol Sci.*, **14 (6)**, 729–739.

Resseguier, V., E. Mémin, and B. Chapron, 2017a: Geophysical flows under location uncertainty, Part I Random transport and general models. *Geophys. & Astro. Fluid Dyn.*, **111 (3)**, 149–176.

Resseguier, V., E. Mémin, and B. Chapron, 2017b: Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading. *Geophys. & Astro. Fluid Dyn.*, **111 (3)**, 177–208, doi:10.1080/03091929.2017.1312101, URL https://hal.inria.fr/hal-01391476.

Resseguier, V., E. Mémin, and B. Chapron, 2017c: Geophysical flows under location uncertainty, Part III SQG and frontal dynamics under strong turbulence conditions. *Geophys. & Astro. Fluid Dyn.*, **111 (3)**, 209–227, doi:10.1080/03091929.2017.1312102, URL https://hal.inria.fr/hal-01391484.

Resseguier, V., E. Mémin, D. Heitz, and B. Chapron, 2017d: Stochastic modelling and diffusion modes for proper orthogonal decomposition models and small-scale flow analysis. *J. Fluid Mech.*, **828**, 29.

Sawford, B., 1986: Generalized random forcing in random-walk models of turbulent dispersion model. *Phys. Fluids*, **29**, 3582–3585.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, **612**, 3079–3012.

Suzuki, N., and B. Fox-Kemper, 2016: Understanding stokes forces in the wave-averaged equations. *J. Geophys. Res. Oceans*, **121**, 3579–3596.

Teixeira, M., and S. Belcher, 2010: On the structure of langmuir turbulence. *Ocean Modell.*, **31 (3)**, 105–119.

Weiss, J., 1991: The dynamics of enstrophy transfer in two-dimensional hydrodynamics. *Physica D*, **48 (2-3)**, 273–294.

Yang, Y., and E. Mémin, 2019: Estimation of physical parameters under location uncertainty using an ensemble^2-expectation-maximization algorithm. *QJRMS*, **145 (719)**, 418–433.