Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading

Valentin Resseguier, Etienne Mémin, Bertrand Chapron

Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading

V. Resseguier†‡**, E. Mémin† and B. Chapron‡
†Inria, Fluminance group, Campus universitaire de Beaulieu, Rennes Cedex 35042, France
‡Ifremer, LOPS, Pointe du Diable, Plouzané 29280, France

November 7, 2016

Abstract

Models under location uncertainty are derived assuming that a component of the velocity is uncorrelated in time. The material derivative is accordingly modified to include an advection correction, inhomogeneous and anisotropic diffusion terms and a multiplicative noise contribution. In this paper, simplified geophysical dynamics are derived from a Boussinesq model under location uncertainty. Invoking usual scaling approximations and a moderate influence of the subgrid terms, stochastic formulations are obtained for the stratified Quasi-Geostrophy (QG) and the Surface Quasi-Geostrophy (SQG) models. Based on numerical simulations, benefits of the proposed stochastic formalism are demonstrated. A single realization of models under location uncertainty can restore small-scale structures. An ensemble of realizations further helps to assess model error prediction and outperforms perturbed deterministic models by one order of magnitude. Such a high uncertainty quantification skill is of primary interests for assimilation ensemble methods. MATLAB® code examples are available online.
Keywords: stochastic sub-grid parameterization, uncertainty quantification, ensemble forecasts.

1 Introduction

Ensemble forecasting and filtering are widely used in geophysical sciences for forecasting and climate projection. In practice, dynamical models are randomized through their initial conditions and a Gaussian error model, and are generally found to be underdispersive (Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013; Berner et al., 2011; Snyder et al., 2015) with a low variance. As a consequence, errors are underestimated and observations are hardly taken into account. Corrections are considered by incorporating inflation procedures or hyperprior to increase the variance of ensemble Kalman filters (Anderson and Anderson, 1999; Bocquet et al., 2015). However, such corrections do not provide an accurate spatial localization of the errors.

Another difficulty of ensemble methods lies in the huge dimensions of the involved state spaces. For obvious computational reasons, ensembles for geophysical applications appear constrained and limited to small sizes. It thus becomes primordial to build strategies to best track the most likely

**Corresponding author. Email: valentin.resseguier@inria.fr

dynamical events. From this point of view, ensemble simulations and stochastic dynamics have clear advantages over the deterministic models.

The simplest random models are defined from Langevin equations with linear damping and additive isotropic Gaussian noise, as, for instance, the linear inverse models (e.g. Penland and Matrosova, 1994; Penland and Sardeshmukh, 1995), or the Eddy-Damped Quasi Normal Markovian (EDQNM) models (e.g. Orszag, 1970; Leith, 1971; Chasnov, 1991). Among other empirical stochastic models, the Stochastic Kinetic Energy Backscatter (SKEBS) (Shutts, 2005; Berner et al., 2009, 2011) and the Stochastic Perturbed Physics Tendency scheme (SPPT) (Buizza et al., 1999) introduce correlated multiplicative noises. SPPT and SKEBS methods have been successfully applied in operational weather forecast centers (Franzke et al., 2015). To target highly non-Gaussian distribution of fluid dynamics properties, an attractive path is to infer randomness from physics (Berner et al., 2015). For this purpose, the time-scale separation assumption is convenient. Hasselmann (1976) already relied on it for geophysical fluid dynamics. This assumption is the foundation of averaging and homogenization theories (Kurtz, 1973; Papanicolaou and Kohler, 1974; Givon et al., 2004; Gottwald and Melbourne, 2013; Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013; Franzke et al., 2015; Gottwald et al., 2015). A successful application of homogenization theory in geophysics is the MTV algorithms (Majda et al., 1999, 2001; Franzke et al., 2005; Majda et al., 2008). The homogenized dynamics is cubic with correlated additive and multiplicative (CAM) noises. This noise structure is able to produce intermittency and extreme events. In practice, the non-linearity of the small-scale equation (fast dynamics) is conveniently replaced by a noise and a damping terms before the homogenization procedure. Noise statistics are estimated from data, with Gaussian assumptions.

In Resseguier et al. (2017a), following Mémin (2014), another approach has been considered to help derive models under location uncertainty based on stochastic calculus and the Ito-Wentzell formula (Kunita, 1997). Mikulevicius and Rozovskii (2004) and Flandoli (2011) already introduce this methodology. Yet, their works mostly focused on pure mathematical aims: existence and uniqueness of SPDE solutions. For our more practical purpose, the large-scale is understood as sub-sampled in time, and the remaining small-scale velocity component is then considered as uncorrelated in time.

Starting with the definition of the revised transport under location uncertainty (section 2), developments are then carried out to derive and analyze the stochastic versions of Quasi-Geostrophy (QG) and Surface Quasi-Geostrophy (SQG) models with a moderate influence of sub-grid terms (section 3). Numerical results highlight the potential of these models under location uncertainty, especially for ensemble forecast (Section 4).

2 Models under location uncertainty

This section briefly outlines main theoretical results discussed in Resseguier et al. (2017a). The velocity is decomposed between a possibly random large-scale component, $\boldsymbol{w}$, and a time-uncorrelated component, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. The latter is Gaussian, correlated in space with possible inhomogeneities and anisotropy. Hereafter, this unresolved velocity component will further be assumed to be solenoidal. To parameterize those spatial correlations, we apply an infinite-dimensional linear operator, $\boldsymbol{\sigma}$, to a d-dimensional space-time white noise[1], $\dot{\boldsymbol{B}}$.

[1]Formally each coefficient of $(t \mapsto \boldsymbol{B}_t)$ is a cylindrical I_d-Wiener process (see Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007) for more information on infinite dimensional Wiener processes and cylindrical I_d-Wiener processes).

In time, the velocity is irregular. The material derivative, D_t, is then changed. In most cases, it coincides with the stochastic transport operator, $\mathbb{D}_t$, defined for every field, Θ, as follows:

$$\mathbb{D}_t\Theta \triangleq \underbrace{\mathrm{d}_t\Theta}_{\substack{\triangleq\ \Theta(\boldsymbol{x},t+\mathrm{d}t)-\Theta(\boldsymbol{x},t)\\ \text{Time increment}}} + \underbrace{(\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\Theta}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)}_{\text{Diffusion}}\mathrm{d}t, \tag{1}$$

where the time increment term $\mathrm{d}_t\Theta$ is used in place of the partial time derivative, as Θ is in general non-differentiable. The diffusion coefficient matrix, $\boldsymbol{a}$, is solely defined by the one-point one-time covariance of the unresolved displacement per unit of time:

$$\boldsymbol{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T = \frac{\mathbb{E}\left\{\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)^T\right\}}{\mathrm{d}t}, \tag{2}$$

and the modified drift is given by

$$\boldsymbol{w}^\star = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T. \tag{3}$$

With this modified material derivative (1), the transport equations under location uncertainty involve three new terms: a modification of the large-scale advection ($\boldsymbol{w}^\star$ instead of $\boldsymbol{w}$), an inhomogeneous and anisotropic diffusion and a multiplicative noise. This random forcing is directly related to the advection by the unresolved velocity.

For incompressible flows ($\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star = 0$), the energy of any tracer, Θ, is conserved for each realization:

$$\mathrm{d}\int_\Omega \Theta^2 = 0, \tag{4}$$

where Ω is the spatial domain. This still holds for active tracers. The diffusion dissipates as much energy as the multiplicative noise is injecting it in the system. In particular, the (ensemble) mean of the energy, $\mathbb{E}\int_\Omega \Theta^2$, is conserved. This results ensures a constant balance between the energy of the mean and the (ensemble) variance. The energy fluxes in these stochastic models are more thoroughly described in Resseguier et al. (2017a).

A random version of the Reynolds transport theorem can further be derived (Mémin, 2014; Resseguier et al., 2017a). From this theorem, usual conservation of mechanics (mass, linear momentum, energy and amount of substance) can be expressed in a stochastic sense. Random Navier-Stokes and Boussinesq models can then be derived. This last model describes the stochastic transports of velocity and density anomaly, as well as incompressibility conditions.

3 Mesoscales under moderate uncertainty

To simplify the stochastic Boussinesq model of Resseguier et al. (2017a), Quasi-Geostrophic (QG) models are developed for large horizontal length scales, L, such as:

$$\frac{1}{Bu} = \left(\frac{Fr}{Ro}\right)^2 = \left(\frac{L}{L_d}\right)^2 \sim 1 \text{ and } \frac{1}{Ro} = \frac{Lf_0}{U} \gg 1, \tag{5}$$

where U is the horizontal velocity scale, $L_d \triangleq \frac{Nh}{f}$ is the Rossby deformation radius, N is the stratification (Brunt-Väisälä frequency) and h is the characteristic vertical length scale. The Rossby deformation radius explicitly defines the mesoscale range, over which both kinetic and buoyancy effects are important, and strongly interact. In the following, both differential operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ, represent 2D operators.

3.1 Specific scaling assumptions

Hereafter, we explicit scaling assumptions to derive the non-dimensional version of the stochastic Boussinesq model.

3.1.1 Quadratic variation scaling

Besides traditional ones, another dimensionless number, Υ, is introduced to relate the large-scale kinetic energy to the energy dissipation due to the horizontal small-scale random component. In the following, $\boldsymbol{\sigma}_{H\bullet}$ stands for the horizontal component of $\boldsymbol{\sigma}$, $\boldsymbol{a}_H$ for $\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T$ and A_u for its scaling. The new dimensionless number is defined by:

$$\Upsilon \triangleq \frac{UL}{A_u} = \frac{U^2}{A_u/T}. \tag{6}$$

This number compares horizontal advective and diffusive terms in the momentum and buoyancy equations. This number can also be related to the ratio between the Mean Kinetic Energy (MKE), U^2, and the Turbulent Kinetic Energy (TKE), A_u/T_σ, where T_σ is the small-scale correlation time. This reads:

$$\Upsilon = \frac{1}{\epsilon}\frac{MKE}{TKE}, \tag{7}$$

where $\epsilon = T_\sigma/T$ is the ratio of the small-scale to the large-scale correlation times. This parameter, ϵ, is central in homogenization and averaging methods (Majda et al., 1999; Givon et al., 2004; Gottwald and Melbourne, 2013). The number Ro/Υ can then be stated to measure the ratio between sub-grid terms and the Coriolis force. In the usual deterministic case and the limit of small Rossby number, the predominant terms of the horizontal momentum equation then correspond to the geostrophic balance. In the stochastic case, this balance also applies from weak ($\Upsilon \gg 1$) to moderate ($\Upsilon \sim 1$) uncertainty. However, if Υ/Ro is close enough to $O(1)$, this geostrophic balance is modified due to the diffusion effects introduced by the small-scale random velocity. Hereafter, developments focus on the moderate uncertainty case. Resseguier et al. (2017b) deals with the strong uncertainty case.

To evaluate Υ for a given flow at a given scale, eddy viscosity or diffusivity values help the determination of A_u. Boccaletti et al. (2007) give some examples of canonical values. Then, the typical resolved velocity and length scale lead to Υ. If no canonical values are known, absolute diffusivity or similar mixing diagnoses could be measured (Keating et al., 2011) as a proxy of the variance tensor.

3.1.2 Vertical unresolved velocity

The scaling to compare vertical to horizontal unresolved velocities is also considered:

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{Ro}{Bu} D, \tag{8}$$

where $D = \frac{h}{L}$ is the aspect ratio and the subscript H indicates horizontal coordinates. This scaling can be derived from the ω-equation (Giordani et al., 2006). For any velocity $\mathfrak{u} = (\mathfrak{u}_H, \mathfrak{w})^T$, which scales as $(\mathfrak{U}, \mathfrak{U}, \mathfrak{W})^T$, this equation reads

$$f_0^2 \partial_z^2 \mathfrak{w} + N^2 \Delta \mathfrak{w} = \boldsymbol{\nabla} \cdot \boldsymbol{Q} \approx -\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathfrak{u}_H^T \boldsymbol{\nabla} \mathfrak{b}) \approx -f_0 \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \mathfrak{u}_H^T \partial_z \mathfrak{u}_H^\perp \right), \tag{9}$$

where $\mathfrak{b}$ stands for the buoyancy variable and $\boldsymbol{Q}$ for the so-called $\boldsymbol{Q}$-vector. In its non-dimentional version, the ω-equation reads:

$$\frac{\mathfrak{W}}{\mathfrak{U}} \left(\partial_z^2 \mathfrak{w} + Bu \Delta \mathfrak{w} \right) \approx DRo \boldsymbol{\nabla} \cdot \boldsymbol{Q}. \tag{10}$$

At planetary scales, Burger number is small and the rotation dominates the stratification, $\frac{\mathfrak{W}}{\mathfrak{U}} \sim DRo$. At smaller scales, with a larger Burger number, the stratification dominates the rotation, $\frac{\mathfrak{W}}{\mathfrak{U}} \sim DRo/Bu$. For the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, the latter is thus more relevant.

Note that the angle between the small-scale component and the horizontal one can be assumed to be constrained by the angle between the isopycnical and the horizontal plane. Invoked to describe baroclinic instabilities theory, this statement helps to specify the anisotropy of the eddy diffusivity (Vallis, 2006). The argument of the orientation of the eddies activity with isentropic surfaces and the related mixing is also supported by several other authors (Gent and Mcwilliams, 1990; Pierrehumbert and Yang, 1993).

In the case of QG models, the large and small Burger scaling cases lead to the same result: the unresolved velocity is mainly horizontal.

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{Ro}{Bu} D \ll D. \tag{11}$$

This is consistent with the assumption of a large stratification, *i.e.* flat isopycnicals, if we admit that the eddies activity appears preferentially along the isentropic surfaces. As a consequence, the terms $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z \partial_z$ scale as $\frac{Ro}{Bu}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H \cdot \boldsymbol{\nabla}$. In the QG approximation, the scaling of the diffusion and effective advection terms including $\boldsymbol{\sigma}_{z\bullet}$ are one to two orders smaller (in power of Ro/Bu) than terms involving $\boldsymbol{\sigma}_{H\bullet}$. For any function ξ, the vertical diffusion $\partial_z(\frac{\boldsymbol{\sigma}_{z\bullet}\boldsymbol{\sigma}_{z\bullet}^T}{2}\partial_z\xi)$ is one order smaller than the horizontal-vertical diffusion term $\boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{z\bullet}^T}{2}\partial_z\xi\right)$ and two orders smaller than the horizontal diffusion term $\boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T}{2}\boldsymbol{\nabla}\xi\right)$.

3.1.3 Beta effect

At mid-latitudes, the related term, given by $\beta \stackrel{\triangle}{=} \partial_y f$, is much smaller than the constant part of the Coriolis frequency. Nevertheless, it can govern a large part of the relative vorticity at large scales. The following scaling is thus chosen (Vallis, 2006):

$$\beta y \sim \boldsymbol{\nabla}^\perp \cdot \boldsymbol{u} \sim \frac{U}{L} = Ro f_0. \tag{12}$$

3.2 Stratified Quasi-Geostrophic model under moderate uncertainty

The moderate uncertainty case corresponds to $\Upsilon \sim 1$. Horizontal advective terms and horizontal sub-grid terms are comparable.

Following similar principles than those used to derive the deterministic stratified QG model (Vallis, 2006), a stochastic QG model can be derived (see Appendix B). This QG solution corresponds to the limit of the Boussinesq solution when the Rossby number goes to zero. The resulting potential vorticity (PV), Q, is then found to be conserved, along the horizontal random flow, up to three source terms:

$$\mathbb{D}_t^H Q \quad = \quad \frac{1}{2}\sum_{i,j\in H} \partial_{ij}^2 \left(\boldsymbol{\nabla}^\perp a_{ij} \cdot \boldsymbol{u}\right) \mathrm{d}t \quad - \quad \frac{1}{2}\boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot (\boldsymbol{a}_H f)\right)^T \mathrm{d}t \quad - \quad \mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right], \quad (13)$$

where the QG PV is:

$$Q \triangleq \Delta\psi + f + \left(\frac{f_0}{N}\right)^2 \partial_z^2 \psi, \quad (14)$$

ψ is the streamfunction, $\boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the $\frac{\pi}{2}$ rotation matrix,

$$\boldsymbol{S} = \frac{1}{2}\left[\boldsymbol{\nabla}\boldsymbol{u}^T + (\boldsymbol{\nabla}\boldsymbol{u}^T)^T\right] \text{ and } \boldsymbol{S}_{\sigma \mathrm{d}B_t} = \frac{1}{2}\left[\boldsymbol{\nabla}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T + (\boldsymbol{\nabla}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T)^T\right] \quad (15)$$

denotes the strain rate tensor of the horizontal resolved and unresolved velocities, $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$, respectively. To interpret the source terms, we rather focus on the material derivative of the PV:

$$D_t^H Q \quad = \quad \boldsymbol{\nabla} \cdot (\alpha \boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot (\boldsymbol{a}_H f)\right)^T \mathrm{d}t - \mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right], \quad (16)$$

with

$$\alpha^T \triangleq \sum_i \left(\boldsymbol{\nabla}\boldsymbol{\sigma}_{H_i}^T\right)^2, \quad (17)$$

which can be decomposed into a symmetric part, positive or negative diffusion of the stream function, and an anti-symmetric part, skew diffusion advection of the stream function. Compared to the traditional QG model, this system includes two smooth (continuous) source/sink terms that depend on the variance tensor, and a random forcing term. The first source term in (16) is correlated in time and may decrease or increase the PV energy. This term is due to the spatial variations of both the diffusion coefficient and the drift correction. The second term takes into account interactions between the Coriolis frequency, including beta effects, and inhomogeneous sub-grid eddies. The last source term in (16) is a noise term, encoding the interactions between the resolved and the unresolved strain rate tensors. Uncorrelated in time, this noise increases the potential enstrophy along time.

To further understand this source term, let us denote Ξ and Λ the eigenvalues associated with the stable directions (*i.e.* negative eigenvalue) of the strain rate tensors of the large-scale flow, $\boldsymbol{S}$, and of the small-scale flow, $\boldsymbol{S}_{\sigma \mathrm{d}B_t}$ respectively. We note θ, the angle between these two stable directions

$$-\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right] \quad = \quad 2\underbrace{\Xi\Lambda}_{>0}\sin(2\theta). \quad (18)$$

The detailed derivation is provided in Appendix B. This random source vanishes when the stable directions of $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$ are aligned or orthogonal. It is maximum and positive (respectively minimum and negative) when there is an angle of $\frac{\pi}{4}$ (respectively $-\frac{\pi}{4}$) between those directions. Around the local position $\boldsymbol{x}$, stable and unstable directions of the large-scale velocity define 2 axes and 4 quadrants. As understood, the strain rate tensor does not depend on the local vorticity. Yet, an hyperbolic deformation will almost resemble a positive vorticity in the upper-left and bottom-right quadrants, and a negative vorticity in the upper-right and bottom-left quadrants. For $\theta = \frac{\pi}{4}$, the stable direction of the small-scale velocity aligns along the upper-left to bottom-right direction. The small-scale velocity then compresses the flow in this direction and dilates the flow in the orthogonal direction (upper-right to bottom-left). The quadrants associated with a seemingly positive (resp. negative) vorticity are brought closer (resp. farther) to $\boldsymbol{x}$. Accordingly, the vorticity increases at $\boldsymbol{x}$. For $-\frac{\pi}{4}$, the vorticity would decrease.

Note the $\mathrm{d}\boldsymbol{B}_t$ factor has been omitted in the right-hand side of equation (18). This term remains a linear function of the uncorrelated noise $\boldsymbol{z} \mapsto \mathrm{d}\boldsymbol{B}_t(\boldsymbol{z})$. Whatever the angle between the stable directions, the source term always has a zero (ensemble) mean and increases the enstrophy since it is a term in $\mathrm{d}\boldsymbol{B}_t$. Equation (18) could then be used to define the horizontal inhomogeneous small-scale component of the velocity. If the conservation of PV is a strong constraint, this component can indeed be defined to ensure that its stable direction is always along or orthogonal to the stable direction of $\boldsymbol{u}$.

A two-layer model could also be deduced from equation (13) or (16). This would help identifying the stochastic parameterization effects on the barotropic and baroclinic modes. In particular, the particular forms of the operator $\boldsymbol{\sigma}$ able to trigger barotropization effects can bemore efficiently studied.

In the stochastic QG model, the stream function ψ is related to the buoyancy, b, the pressure, p', and the velocity, $\boldsymbol{u}$, by the usual relations:

$$b = f_0 \partial_z \psi,\ p' = \rho_b f_0 \psi \text{ and } \boldsymbol{u} = \boldsymbol{\nabla}^\perp \psi, \tag{19}$$

where ρ_b is the mean (background) density. The horizontal noise term, $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$, appearing in both the horizontal stochastic material derivative and in the 2×2 horizontal variance tensor, $\boldsymbol{a}_H$, is in geostrophic balance with a pressure component uncorrelated in time. Due to their scaling, the vertical noise and its variance are neglected in the final equations.
For homogeneous turbulence conditions, the transport of PV (16) simplifies. The variance tensor becomes constant, the first two source terms disappear, to give

$$\mathbb{D}_t^H Q = D_t^H Q = -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\boldsymbol{\sigma}\mathrm{d}B_t}\right]. \tag{20}$$

The transport of the PV (equation (13) or (20)) determines the dynamics of the fluid interior. Boundary conditions are then necessary to specify completely the dynamics.

3.3 Surface Quasi-Geostrophic model under moderate uncertainty

A classical choice considers a vanishing solution in the deep ocean and a buoyancy transport at the surface (Vallis, 2006; Lapeyre and Klein, 2006):

$$\psi \underset{z\to-\infty}{\longrightarrow} 0 \text{ and } D_t^H b_{|_{z=0}} = \mathbb{D}_t^H b_{|_{z=0}} = 0. \tag{21}$$

Assuming zero PV in the interior but keeping these boundary conditions leads to the Surface Quasi-Geostrophic model (SQG) (Blumen, 1978; Held et al., 1995; Lapeyre and Klein, 2006;

Constantin et al., 1994, 1999, 2012). Under the stochastic framework, the derivation is similar. The PV is indeed identical to the classical one (see equation (14)), assuming zero PV in the interior and vanishing solution as $z \to -\infty$ unsurprisingly yields the same SQG relationship:

$$\hat{b} = N\|\boldsymbol{k}\| \, \hat{\psi}. \tag{22}$$

The top boundary condition, equation (21), provides an evolution equation, namely the horizontal transport of surface buoyancy, in the stochastic sense:

$$\mathbb{D}_t^H b = 0. \tag{23}$$

The time-uncorrelated component of the velocity, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, is divergence-free. Its inhomogeneous and anisotropic spatial covariance has then to be specified. The time-correlated component of the velocity is also divergence-free, with a stream function specified by the SQG relation (22). The buoyancy is randomly advected, and the resulting smooth velocity component is random as well.

3.4 Summary

For simplified models, stochastic versions are derived for scaling assumptions related to the sub-grid terms. For moderate uncertainty, the PV is transported along the random flow up to three source terms. The first one, smooth in time, is due to spatial variations of the inhomogeneous diffusion and the drift correction. The second one, also smooth, encodes the interaction between inhomogeneous turbulence and Coriolis frequency. These terms disappear for an homogeneous turbulence. The last term, a time-uncorrelated multiplicative noise, involves the large-scale and the small-scale strain rate tensors. It is a source of potential enstrophy and its instantaneous value depends on the angle between the large-scale and small-scale stable directions. Assuming zero PV in the interior, a SQG model follows from this QG model.

4 Numerical results

We focus on this SQG_{MU} model (3.3). A high-resolution deterministic SQG simulation provides a reference. The MATLAB® codes are available online (`http://vressegu.github.io/sqgmu`). Numerical results are analyzed in terms of the resolution gains (when a single realization is simulated) and the potential for ensemble forecasting in estimating spatial and spectral reconstruction errors (for an ensemble of realizations).

4.1 Test flow

The initial conditions defining the test flow, Figure 1, consist of a spatially smooth buoyancy field with two warm elliptical anticyclones and two cold elliptical cyclones given by:

$$\begin{aligned} b(\boldsymbol{x}, t=0) \quad = \quad & F\left(\boldsymbol{x} - \begin{pmatrix} 250 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right) + F\left(\boldsymbol{x} - \begin{pmatrix} 750 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right) \\ & -F\left(\boldsymbol{x} - \begin{pmatrix} 250 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right) - F\left(\boldsymbol{x} - \begin{pmatrix} 750 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right), \end{aligned} \tag{24}$$

with

$$F(\boldsymbol{x}) \triangleq B_0 \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right) \text{ and } \left\{\begin{array}{lcl}\sigma_x &=& 67 \text{ km},\\ \sigma_y &=& 133 \text{ km}.\end{array}\right. \tag{25}$$

The size of the vortices is of order of the Rossby radius L_d. The buoyancy and the stratification have been set with: $B_0 = 10^{-3} m.s^{-2}$ and $N = 3f_0$. The Coriolis frequency is set to $1.028 \times 10^{-4} s^{-1}$ (45° N). Periodic boundaries conditions are considered.

The deterministic high-resolution SQG reference model is associated with a spatial mesh grid of 512^2 points, whereas the low-resolution (deterministic or stochastic) SQG models are run on 128^2 points. The simulations have been performed through a pseudo-spectral code in space. As for the temporal discrete scheme the deterministic simulation relies on a fourth-order Runge-Kutta scheme, whereas the stochastic ones are based on an Euler-MaruyamaEuler-Maruyama scheme (Kloeden and Platen, 1999). For our application, the weak precision of this scheme is balanced by the use of a small time step. In all the simulations (deterministic and random, high-resolution and low-resolution), a standard hyperviscosity model is used:

$$\mathbb{D}_t b = \alpha^{hv} \Delta^4 b \, \mathrm{d}t, \tag{26}$$

with a coefficient $\alpha^{hv} = (5 \times 10^{29} m^8.s^{-1}) M_x^{-8}$ where M_x denotes the meshgrid size (*i.e.* 128 or 512).

Figure 1 displays the high-resolution buoyancy field at $t = 0, 5, 10, 13, 15, 16, 20$ and 30 days. During the first ten days, the vortices turn with slight deformation. Vortices of the same sign have their tails that draw closer. This creates high shears around four saddle points located at $(x, y) = (0, 250), (500, 250), (0, 750)$ and $(500, 750)$ (in km). A strong non-linearity in the neighborhood of a saddle point has been identified to become a major source of instability (Constantin et al., 1994, 1999, 2012). In our case, this effect is weak but yields an effective creation of turbulence 10 days later. Shears create long and fine filaments, wrapping around the vortices until the 15^{th} day. At this time, the filaments become unstable, break and a so-called "pearl-necklace" appears, characteristic of the SQG model, days 17-18 in the simulation. These small vortices are then ejected from their orbits. Between days 17^{th} and 25^{th}, they interact with the large vortices, the filaments and other small vortices, to create a fully-developed SQG turbulence orbiting around the four large vortices.

4.2 Simulation of the random velocity

To simulate the $\mathrm{SQG_{MU}}$ model (22-23), the covariance of the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ must be specified. As this unresolved velocity field is assumed divergence-free, we introduce the following stream function linear operator, $\breve{\psi}_\sigma$, and its kernel, $\breve{\psi}_\sigma$:

$$\boldsymbol{\sigma}_H(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp \psi_\sigma(\boldsymbol{x})\mathrm{d}B_t, = \int_\Omega \mathrm{d}\boldsymbol{z}\ \boldsymbol{\nabla}^\perp_{\boldsymbol{x}} \breve{\psi}_\sigma(\boldsymbol{x}, \boldsymbol{z})\mathrm{d}B_t(\boldsymbol{z}). \tag{27}$$

As such, a single cylindrical Wiener process, B_t, is sufficient to sample our Gaussian process. This is specific to two-dimensional domains. In 3D, a vector of 3 independent $\mathbb{I}_d$-cylindrical Wiener processes, and a projection operator on the divergence-free vector space or a curl must be considered to simulate an isotropic small-scale velocity (Mémin, 2014). For a divergent unresolved velocity, equation (27) can additionally involve the gradient of a random potential, $\boldsymbol{\nabla}\tilde{\psi}_\sigma \mathrm{d}B_t$.

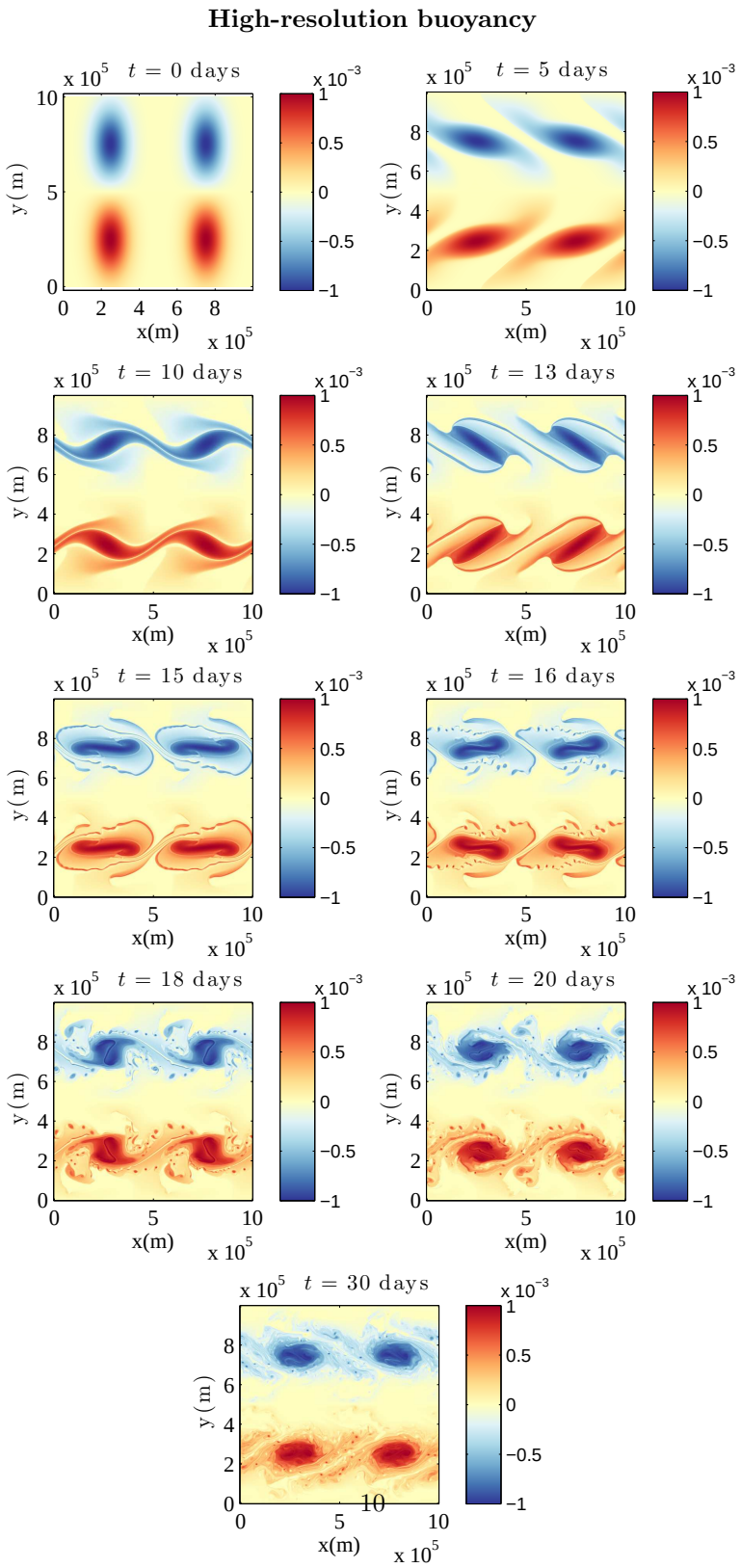

Figure 1: Buoyancy $(m.s^{-2})$ at $t = 0, 5, 10, 13, 15, 16, 20, 30$ days of advection for the usual SQG model at resolution 512^2.

Then, similar to the Kraichnan's model, a solenoidal homogeneous field can be considered: (Kraichnan, 1968, 1994; Gawędzki and Kupiainen, 1995; Majda and Kramer, 1999):

$$\boldsymbol{\sigma}_H(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \int_\Omega \mathrm{d}\boldsymbol{z}\ \boldsymbol{\nabla}^\perp_{\boldsymbol{x}}\breve{\psi}_\sigma(\boldsymbol{x}-\boldsymbol{z})\mathrm{d}B_t(\boldsymbol{z}) = \left(\boldsymbol{\nabla}^\perp\breve{\psi}_\sigma \star \mathrm{d}B_t\right)(\boldsymbol{x}). \tag{28}$$

where $\star$ denotes a convolution. Although spatially inhomogeneous field would be more physically relevant, homogeneity greatly simplifies the random field simulation. Indeed, homogeneity in physical space implies independence between the Fourier modes

$$\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k}) = i\boldsymbol{k}^\perp\widehat{\breve{\psi}_\sigma}(\boldsymbol{k})\hat{\dot{B}}(\boldsymbol{k}), \tag{29}$$

in the half-space $\boldsymbol{k} \in (\mathbb{R}\times\mathbb{R}^{+*}) \cup (\mathbb{R}^+\times\{0\})$. Thus, the small-scale velocity can be conveniently specified from its omnidirectional spectrum:

$$\boldsymbol{k} \mapsto \frac{1}{\mu(\Omega)}\mathbb{E}\oint_{[0,2\pi]}\mathrm{d}\theta_{\boldsymbol{k}}\|\boldsymbol{k}\|\left\|\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k})\right\|^2 = \frac{2\pi}{\Delta t}\|\boldsymbol{k}\|^3\left|\widehat{\breve{\psi}_\sigma}\left(\|\boldsymbol{k}\|\right)\right|^2, \tag{30}$$

where $\mu(\Omega)$ is the surface of the spatial domain Ω, $\theta_{\boldsymbol{k}}$ is the angle of the wave-vector $\boldsymbol{k}$ and Δt the simulation time-step. Consistent with SQG turbulence, the omni-directional spectrum slope, denoted s, is fixed to $-\frac{5}{3}$. For 2D Euler equations, the slope would be set to -3. If the small scales spectrum slope is unknown, the spectrum slope of the resolve scales – estimated on line – may enable to specify s through a scale similarity assumption. The unresolved velocity should be energetic only where the dynamics cannot be properly resolved. Consequently, we apply to the spectrum a smooth band-pass filter, f_{BP}, which has non-zero values between two wavenumbers κ_{min} and κ_{max}. The parameter κ_{min} is inversely related to the spatial correlation length of the unresolved component. In practice, we set κ_{max} to the theoretical resolution, $\frac{\pi}{\Delta x}$, and κ_{min} to the effective resolution (hereafter $\kappa_{min} = \kappa_{max}/2$). Figure 2 illustrates this spectrum specification. The small scales' energy is specified by the diffusion coefficient a_H and the simulation time step:

$$\mathbb{E}\left(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right)\left(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right)^T = \frac{1}{\Delta t}\boldsymbol{a}_H = \frac{1}{\Delta t}\begin{pmatrix} a_H & 0 \\ 0 & a_H\end{pmatrix}. \tag{31}$$

The diagonal structure of the variance tensor is due both to incompressiblity and isotropy. The scalar variance tensor, a_H, is similar to an eddy viscosity coefficient. So, a typical value of eddy viscosity used in practice is a good proxy to setup this parameter. Otherwise, this parameter can be tuned. For this paper, it is set to 9 $m^2.s^{-1}$. The time step depends itself, through the CFL conditions, on both the spatial resolution and the maximum magnitude of the resolved velocity. Finally, equation (29) writes:

$$\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}}\ i\boldsymbol{k}^\perp f_{BP}\left(\|\boldsymbol{k}\|\right)\|\boldsymbol{k}\|^{-\alpha}\frac{\widehat{\mathrm{d}B_t}}{\sqrt{\Delta t}}(\boldsymbol{k}) \text{ with } s = 3-2\alpha = -\frac{5}{3}, \tag{32}$$

where A is a constant to ensure $\mathbb{E}\left\|\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right\|^2 = 2\frac{a_H}{\Delta t}$ (see equation (31) above), $\widehat{\mathrm{d}B}_t$ is the spatial Fourier transform of $\mathrm{d}B_t$, with $\frac{\mathrm{d}B_t}{\sqrt{\Delta t}}$, a discrete scalar white noise process of unit variance in space and time. To sample the small-scale velocity, we first sample $\frac{\mathrm{d}B_t}{\sqrt{\Delta t}}$, to get $\widehat{\frac{\mathrm{d}B_t}{\sqrt{\Delta t}}}$, and finally $\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k})$ with the above equation.

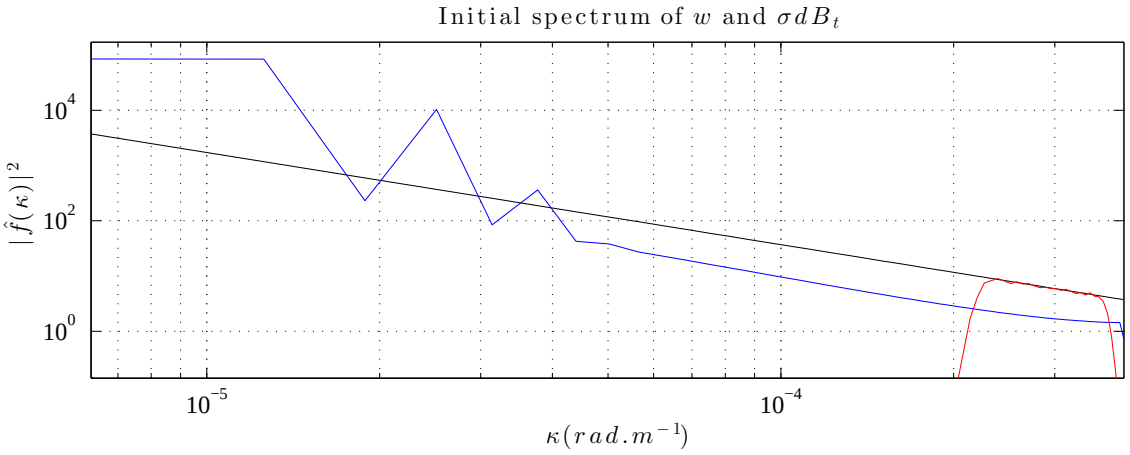

Figure 2: Spectrum ($m^2.s^{-2}$/(rad.m^{-1})) of $\boldsymbol{w}$, at the initial time, in blue, spectrum of $\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}$ (up to a multiplicative constant), in red, and slope $-\frac{5}{3}$ in black. In the simulation performed, $\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}$ is restricted to a narrow spectral band. Thus, this velocity component almost only acts near the resolution cutoff, where the large-scale component, $\boldsymbol{w}$, has a low energy.

4.3 Resolution gain on a single simulation

In Figure 3, the buoyancy field and its spectrum for low resolution SQG_{MU} and deterministic SQG simulations are displayed for the day 17^{th}. For the spectrum plots (right column), the slope $-\frac{5}{3}$ is superimposed. While the spectrum tail of the SQG model falls slightly before the stochastic one, the most significant gain is observed in the spatial domain, i.e. in the phase of the tracer. Indeed, the SQG_{MU} buoyancy field exhibits pearl-necklaces, only obtained at higher resolution. The low-resolved SQG simulation only generates smooth and stable filaments. Though small-scale energy distribution remains similar for both low-resolved models, the phase of the stochastic tracer is more accurate. This may seem surprising since the unresolved velocity, $\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}$, is defined in a loose way, through its spectrum, without prescribing the nature of its phase. However, the noise is multiplicative, and the random forcing, $-(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}})\cdot\boldsymbol{\nabla}b$, does implicitly take into account the tracer phase.

Note, within the stochastic framework, the diffusion coefficient is explicitly related to the noise variance. If the small-scale velocity is set to a magnitude three times smaller than the one prescribed by the diffusion coefficient $\frac{a_H}{2}$, the tracer field becomes quickly too smooth (see Figure 4). Conversely, if the small-scale velocity is set to a magnitude three times larger than dictated by the stochastic transport model, the tracer field becomes rapidly too noisy. This is visible both in the spatial and Fourier spaces (Figure 4). The stochastic transport model thus imposes a correct balance between noise and diffusion.

4.4 Ensemble forecasts

While single realization of SQG_{MU} model carries more valuable information than a deterministic SQG formulation at the same resolution, our model further enables to perform ensemble forecasting

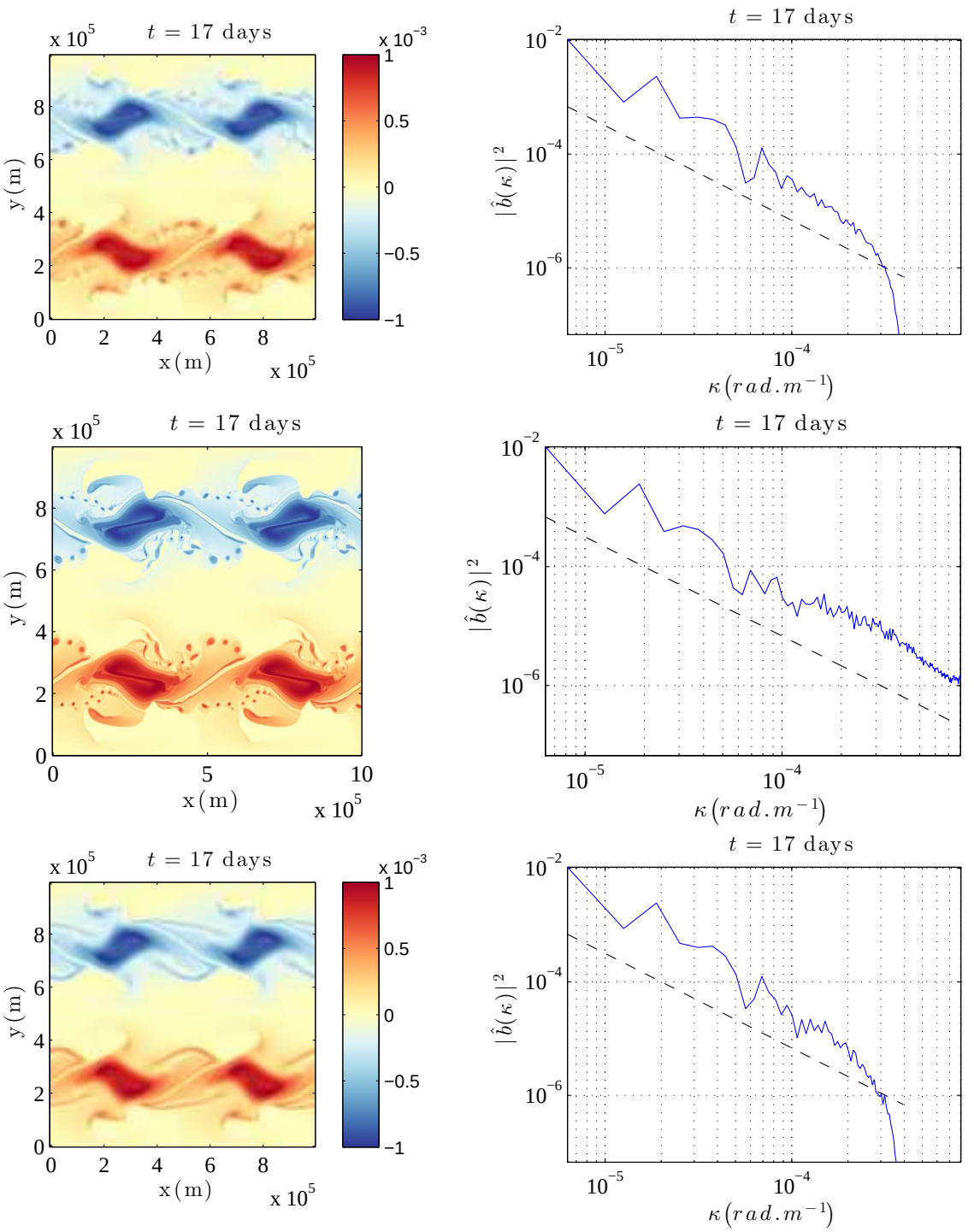

Figure 3: Buoyancy ($m.s^{-2}$) and its spectrum ($m^3.s^{-4}/(\text{rad}.m^{-1})$) at the 17^{th} day of advection for SQG_{MU} at resolution 128^2 (top), SQG at resolution 512^2 (middle) and at resolution 128^2 (bottom). Unlike SQG_{MU}, the low-resolved SQG simulation diffuses the "pearl necklaces", noticeable only at higher resolution.

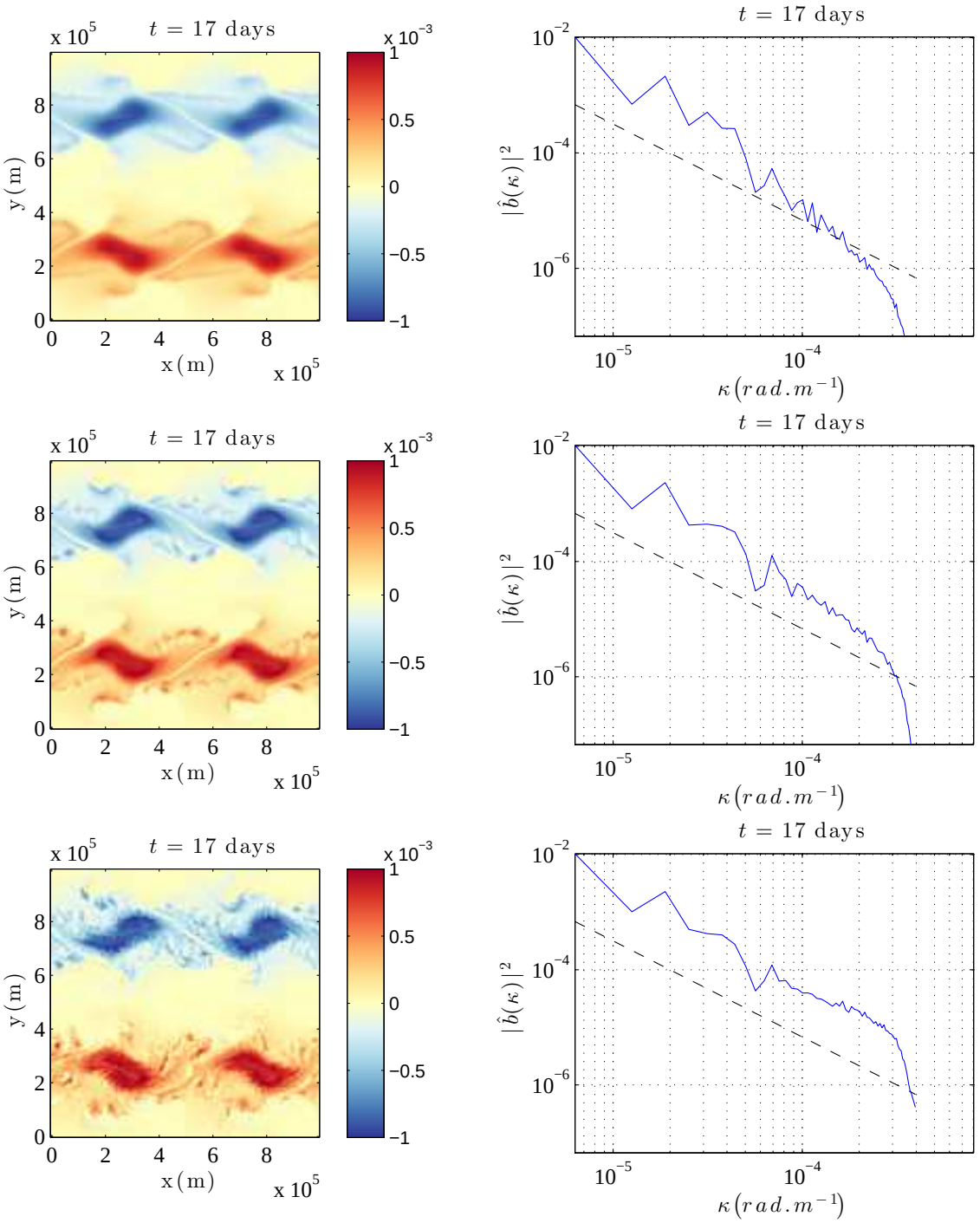

Figure 4: Buoyancy ($m.s^{-2}$) and its spectrum ($m^2.s^{-4}/(\mathrm{rad}.m^{-1})$) at the $t = 17^{th}$ day of advection for the $\mathrm{SQG_{MU}}$ model with a small-scale velocity component three times weaker than the one prescribed by the diffusion coefficient a_H (top), with the correct amount of small-scale energy (middle) and a small-scale velocity three times higher than the model diffusion. If the prescribed balance between noise and diffusion is not met the tracer field becomes quickly too smooth or too noisy.

and filtering. Straightforwardly, an ensemble of independently randomly forced realizations $\{b^{(i)}|i = 1, \cdots, N_e\}$ of tracer b can be simulated according to the SPDE (23). The probability density function and all the statistical moments of the simulated tracer can then be approximated. For instance, the (ensemble) mean of the buoyancy is a spatio-temporal field defined by:

$$\mathbb{E}(b)(\boldsymbol{x}, t) \approx \hat{\mathbb{E}}(b)(\boldsymbol{x}, t) \triangleq \frac{1}{N_e} \sum_{i=1}^{N_e} b^{(i)}(\boldsymbol{x}, t), \tag{33}$$

where N_e denotes the ensemble size. This is in essence a Monte-Carlo Makov Chain (MCMC) simulation. The ensemble size is deliberately kept small[2] in order to assess the proposed stochastic framework skills.

We compare the ensemble bias with the estimated error provided by the ensemble itself. The bias corresponds to the discrepancy between the tracer ensemble mean and the SQG simulation at high resolution[3] (512^2).

Our reference is deterministic since the initial condition is perfectly known and the target dynamics is deterministic, as the real ocean dynamics. The partial knowledge of initial conditions is a complementary issue not addressed in this paper. The reference being deterministic, the bias represents both the error of the mean and the mean of the error:

$$\widehat{\mathbb{E}}\{b\} - b^{ref} = \widehat{\mathbb{E}}\{\epsilon\}, \tag{34}$$

where $\epsilon = b - b^{ref}$ stands for the (random) error. We denote by e the absolute value of this bias. Another error metric could be the Root Mean Square Error (RMSE), $\sqrt{\widehat{\mathbb{E}}\{\epsilon^2\}}$. Yet slightly larger, it is found to have similar spatial and spectral distributions (not shown).

The estimated error, denoted ϵ_{est}, is set to 1.96 times the ensemble standard deviation. This specific value corresponds to the (Gaussian) 95% confidence interval. Although the tracer distribution is not Gaussian, this value provides an accurate conventional error estimate:

$$\epsilon^2_{est}(\boldsymbol{x}, t) = (1.96)^2 \widehat{Var}(b) \triangleq (1.96)^2 \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(b^{(i)} - \hat{\mathbb{E}}(b) \right)^2. \tag{35}$$

As this error depends on time and space, several comparisons are performed at several distinct times in both the spatial and Fourier domains. In Figures 5 and 6, the absolute value of spatial fields (34) and (35) (*i.e.* e and ϵ_{est}) are compared at days 10, 13, 15, 17, 20 and 25. As obtained, the SQG_{MU} model enables the ensemble to predict the positions and the amplitudes of its own errors with a very good accuracy.

To compare the spread-error consistency of the proposed model, a more classical type of random simulation is considered. An ensemble of the same size is initialized with random perturbations of the initial conditions (24). The perturbations are assumed to be homogeneous, isotropic, Gaussian and are sampled from a $(-\frac{5}{3})$ spectrum restricted to the small spatial scales, as shown in Figure 7. Then, the ensemble is forecast with the deterministic SQG model.

Figures 8 and 9 represent the spectrum of the errors. The blue and red lines with crosses stand for the spectrum of the bias absolute value, e, of the SQG_{MU} with deterministic initial conditions and of the SQG model with random initial conditions, respectively.

[2] All the random simulations are performed with 200 – 128^2 mesh-size – realizations.

[3] Note this simulation is afterward spatially filtered and subsampled to the same resolution as the ensemble

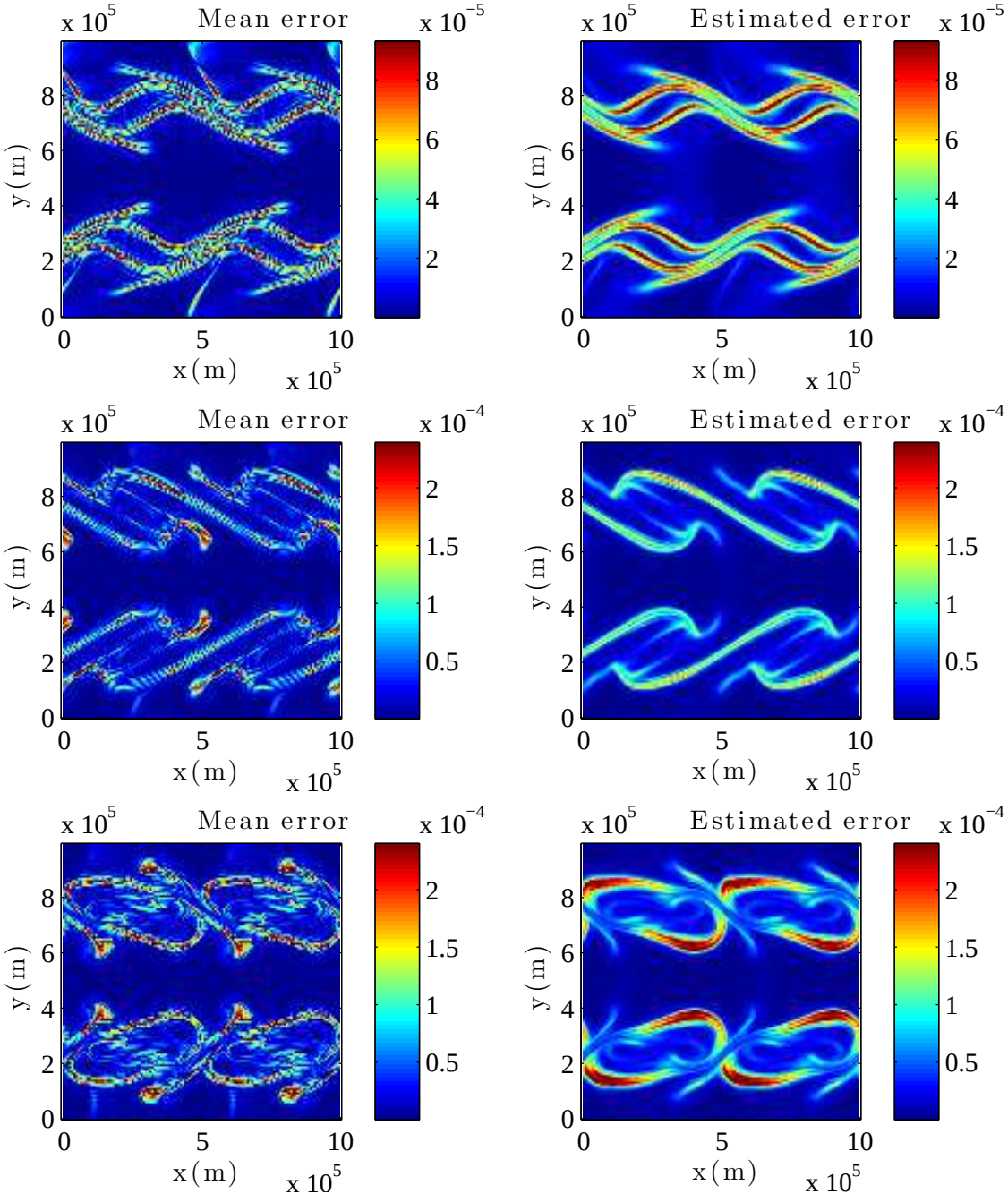

Figure 5: Buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, $(m.s^{-2})$ of the $\mathrm{SQG}_{\mathrm{MU}}$ model (left) and its estimation, ϵ_{est}, (1.96× the standard deviation of the ensemble) (right) at resolution 128^2 at (from top to bottom) $t = 10, 13$ and 15 days of advection. The reference is the usual SQG model at resolution 512^2 – adequately filtered and subsampled.

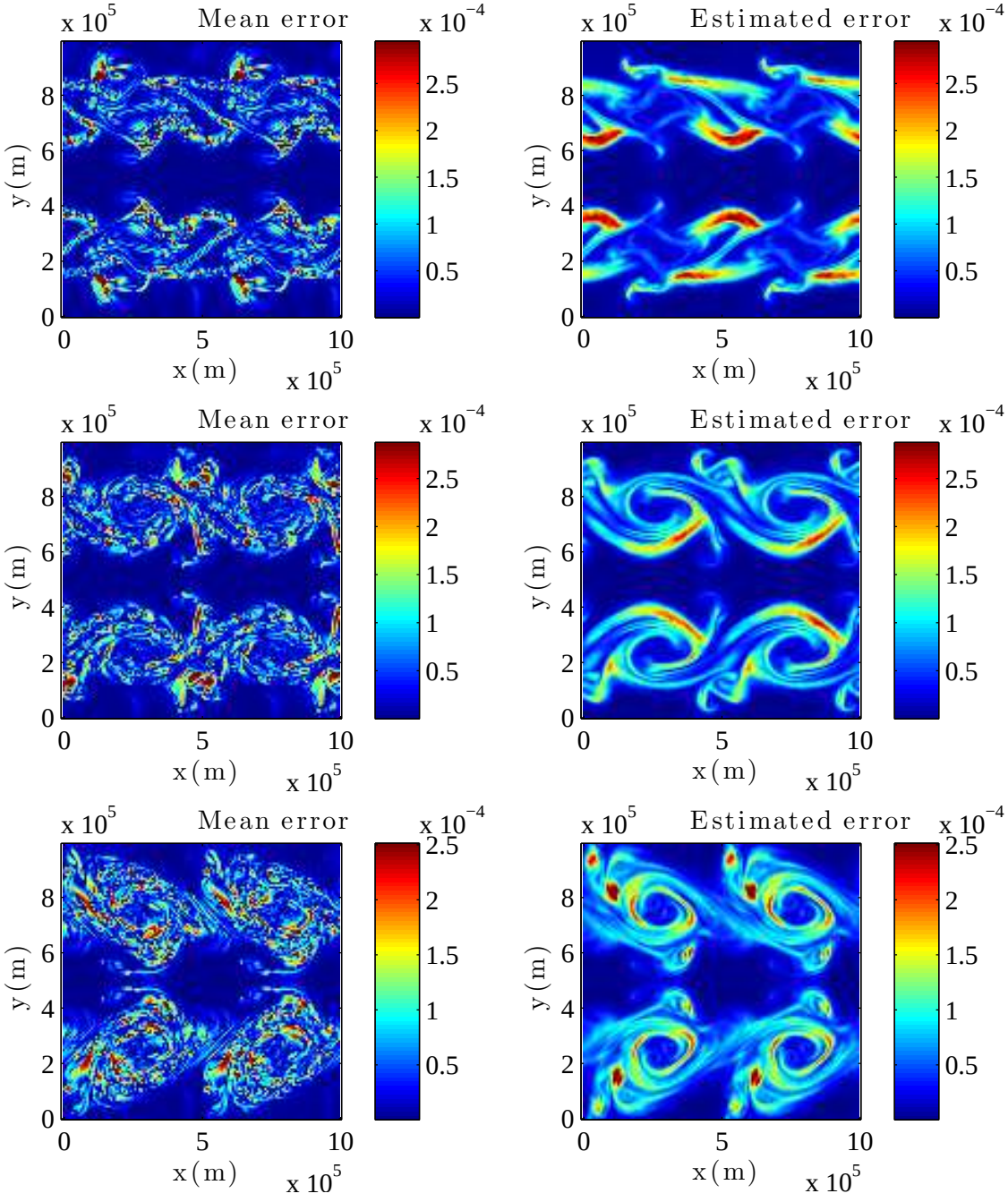

Figure 6: Buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, $(m.s^{-2})$ of the $\mathrm{SQG}_{\mathrm{MU}}$ model (left) and its estimation, ϵ_{est}, ($1.96\times$ the standard deviation of the ensemble) (right) at resolution 128^2 at (from top to bottom) $t = 17, 20$ and 25 days of advection. The reference is the usual SQG model at resolution 512^2 – adequately filtered and subsampled.

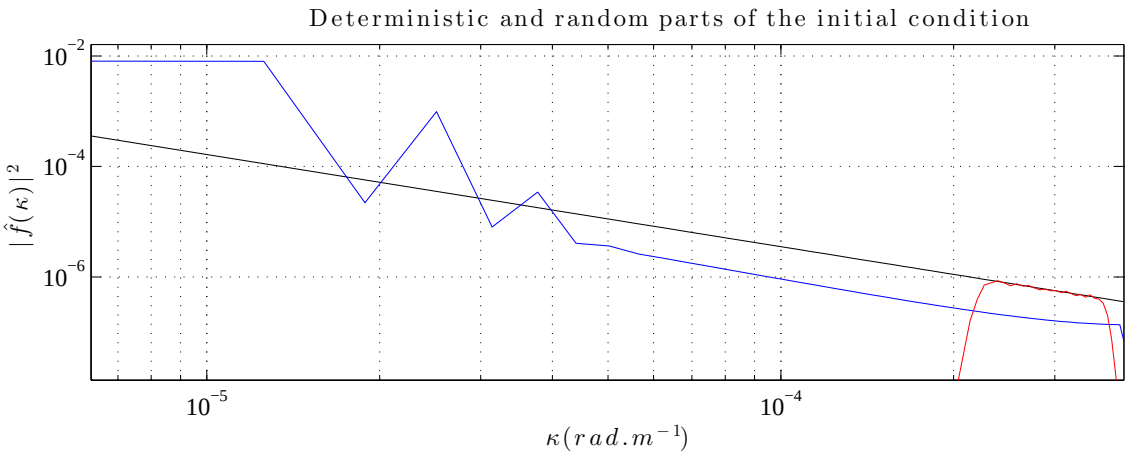

Figure 7: Spectrum ($m^2.s^{-4}/(\text{rad}.m^{-1})$), at the initial time, of the mean buoyancy, in blue, spectrum of its random perturbation, in red, and slope $-\frac{5}{3}$ in black. The initial perturbation is restricted to a narrow spectral band. This random initial condition has been used to simulate an ensemble with the deterministic SQG model.

Deterministic and stochastic models have close distribution of errors over the scales, although the SQG_{MU} ensemble mean generally leads to lower errors than the SQG ensemble mean.

The blue line with circles denotes the spectrum of the SQG_{MU} ensemble estimated error, ϵ_{est}. As a benchmark, we superimposed the spectrum of the same estimator, ϵ_{est}, but simulated with the usual model (red curve with circles). This estimation is dramatically underestimated. It is generally one order of magnitude smaller that the real error. To reduce this drawback, a solution would be to multiply by 10 the perturbations of the initial condition. However, this solution introduces strong errors on the realizations (not shown). Their small-scale errors are generally one order of magnitude larger than the ones of our model. These realizations of the deterministic model remain far from the reference for about ten days. On the contrary, the SQG_{MU} predicts the correct spectral distribution of errors at each time, except at very small-scales, and each of its realizations are accurate as shown in the previous subsection. Let us note however that most of the errors are concentrated at large scales.

SQG_{MU} thus appears to provide a relevant ensemble of realizations, as it enables us to estimate the amplitude of its own error with a good accuracy both in the spatial and spectral domains.

With such an ensemble of realizations, it is now possible to analyze the spatio-temporal evolution of the statistical moments. In Figure 10, we plotted the ensemble tracer mean and variance for $t = 17, 20$ and 30 days of advection. As expected, the mean field is more smooth than the realizations (see Figure 4 for comparison at $t = 17$ days). One realization provides a more realistic field than the mean from a topological point of view. Indeed, the realization exhibits physically relevant small-scale structures. Nevertheless, those structures have uncertain shapes and positions. Therefore, on average, the mean field is closer (in the sense of the norm $\| \bullet \|^2_{L^2(\Omega)}$) to the reference. Besides, those uncertain small-scale structures, forgotten by the mean field, are visible in the variance. The variance becomes significant after 10 days of advection, near the stretched saddle points. The

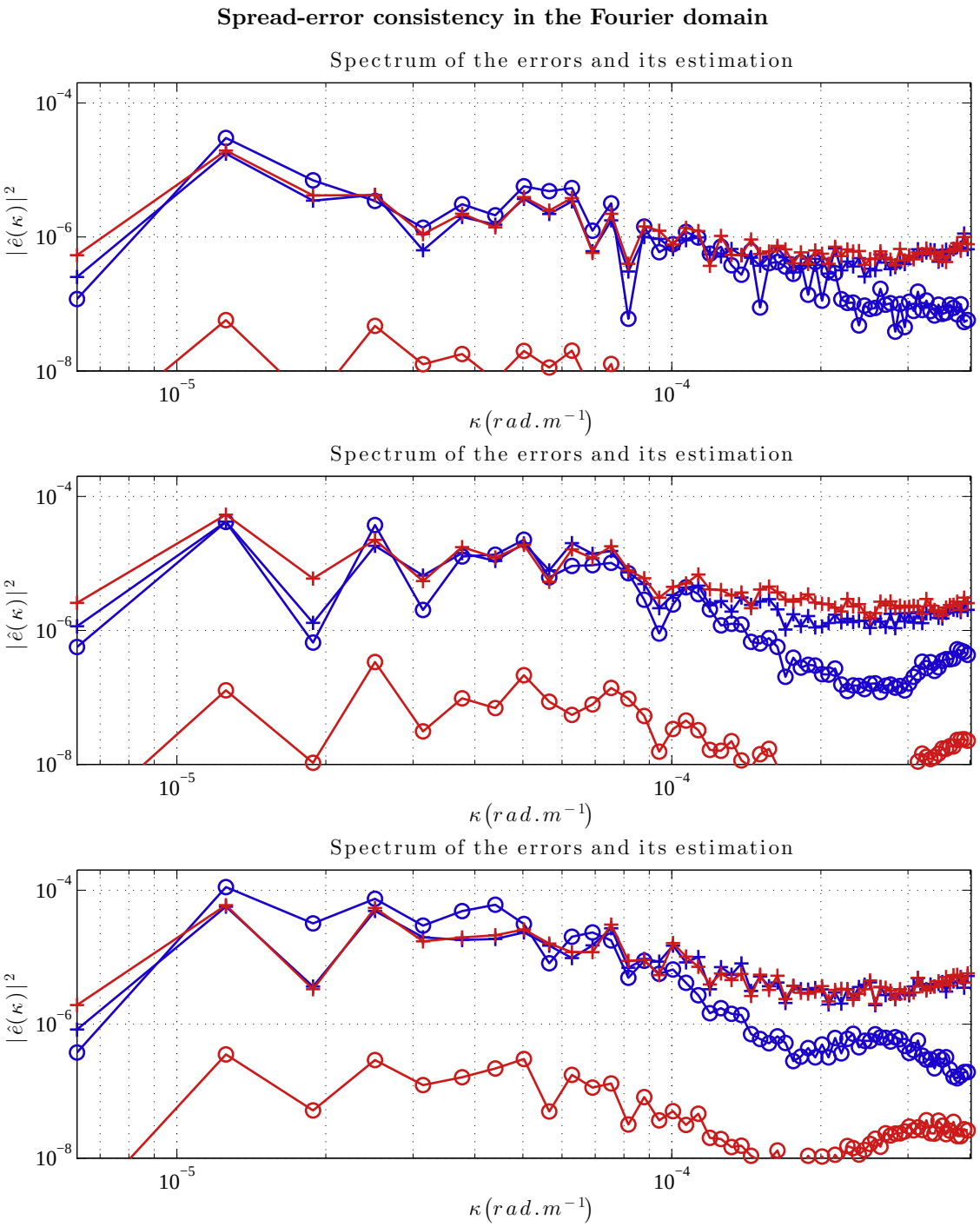

Figure 8: Spectrum of the buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, (lines with crosses) and spectrum of the estimated error, ϵ_{est}, (1.96× the standard deviation of the ensemble) (lines with circles) ($m^2.s^{-4}/(\text{rad}.m^{-1})$) of the low-resolution SQG model with random initial conditions (red) and of the SQG_{MU} model at the same resolution (blue), at (from top to bottom) $t = 10, 13$ and 15 days of advection. The reference is the usual SQG model at resolution 512^2– adequately filtered and subsampled. The low-resolution deterministic model with random initial conditions underestimates the error by at least one order of magnitude whereas our estimation is very precise except at small scales.

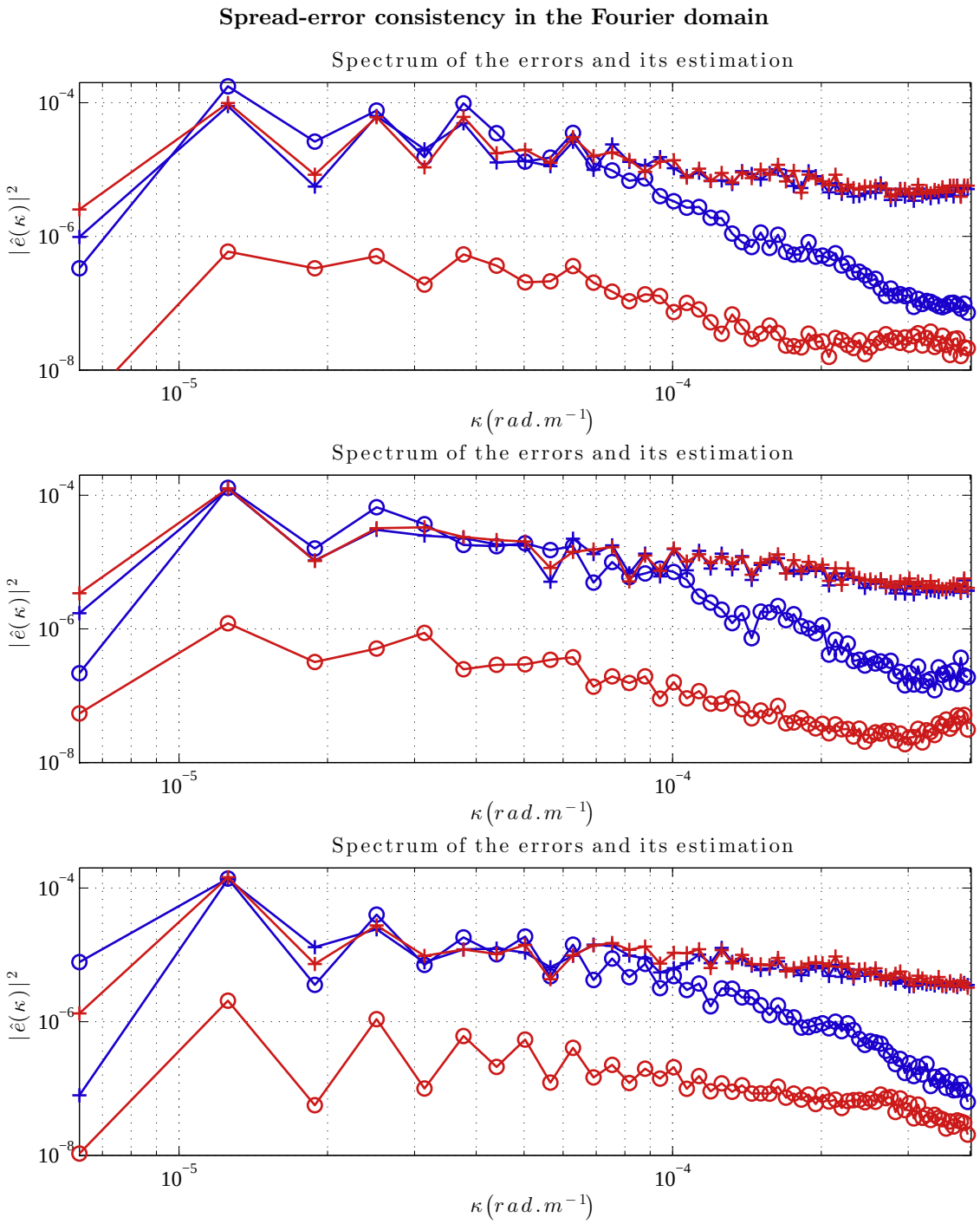

Figure 9: Spectrum of the buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, (lines with crosses) and spectrum of the estimated error, ϵ_{est}, (1.96× the standard deviation of the ensemble) (lines with circles) ($m^2.s^{-4}/(\text{rad}.m^{-1})$) of the low-resolution SQG model with random initial conditions (red) and of the SQG_{MU} model at the same resolution (blue), at (from top to bottom) $t = 17, 20$ and 25 days of advection. The reference is the usual SQG model at resolution 512^2– adequately filtered and subsampled. The low-resolution deterministic model with random initial conditions underestimates the error by at least one order of magnitude whereas our estimation is very precise except at small scales.

strong tracer gradients create strong multiplicative noises. Indeed, strong large-scale gradients involve smaller scales, and thus interact with the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. Then, at $t = 17$ days, the filament instabilities are triggered by the unresolved velocity stretching effects. The appearance of "pearl necklaces" and the underlying motions of those small-scale eddies are mainly determined by the action of the unresolved velocity component. In consequence, these structures are associated with a high uncertainty in their shapes and locations. Hence, they appear naturally on the variance field. At $t = 20$, those sources of variance remain and mushroom-like structures also develop near $(x, y) = (0, 100), (500, 100), (0, 900)$ and $(500, 900)$ (in km). The evolution of these fronts are uncertain, and also show up in the variance field. On the day 30^{th}, these random structures are transported by the zonal jets which are located at $y = 0$ and $y = 500$ km.

The empirical moments of order 3 and 4 can also be evaluated with the ensemble. A high 4^{th} order moment directly relates to the occurrence of extreme events, which is very relevant for dynamical analysis. The point-wise 4-th order moment is centered and normalized to obtain the so-called kurtosis:

$$m_4 \triangleq \frac{\mathbb{E}\left(b - \mathbb{E}(b)\right)^4}{\left(\mathbb{E}\left(b - \mathbb{E}(b)\right)^2\right)^2}. \tag{36}$$

The excess kurtosis, $m_4 - 3$ highlights deviations from Gaussianity. In particular, positive values figure the existence of fat-tail distribution. On the right column of Figure 11, the logarithm of the excess kurtosis is displayed for several distinct times. Negative values of the excess kurtosis (which indicates a flatter peak around the mean) have been set to zero. The "pearl necklaces", identified in the variance plots, engender fat-tailed distribution at days $t = 17$ and 20. The small eddies of a "pearl necklace" have similar vorticity and are close to each other, creating high shears between them. A given eddy can be ejected from the necklace by its closest neighbors, and led up to the north or south down. In such a case, the eddy reaches a zone of the space, neither warm nor cold, with weak variability (e.g. with both local mean and variance being low compared to eddy's temperature). This brings extreme tracer values in statistical homogeneous areas. Finally, the random structures, associated with extreme events are trapped in the zonal jets.

The point-wise moment of order 3 marks the asymmetry of the point-wise tracer distribution. The skewness is the third-order moment of the centered and normalized tracer:

$$m_3 \triangleq \frac{\mathbb{E}\left(b - \mathbb{E}(b)\right)^3}{\left(\mathbb{E}\left(b - \mathbb{E}(b)\right)^2\right)^{\frac{3}{2}}}. \tag{37}$$

Considering the interpretation of excess-kurtosis, the skewness identifies the predominant occurrence of cold (resp. warm) extreme events, associated with the cold (resp. warm) "pearl-necklaces".

5 Conclusion

Models under location uncertainty involve a velocity partially time-uncorrelated. Accordingly, the material derivative, the interpretation of conservation laws, and the usual fluid dynamics models are modified. In this paper, the random Boussinesq model is approximated by the so-called QG equations. In our random framework, the approximation depends on sub-grid terms scaling. With moderate turbulent dissipation, the PV is randomly transported in the fluid interior up to three

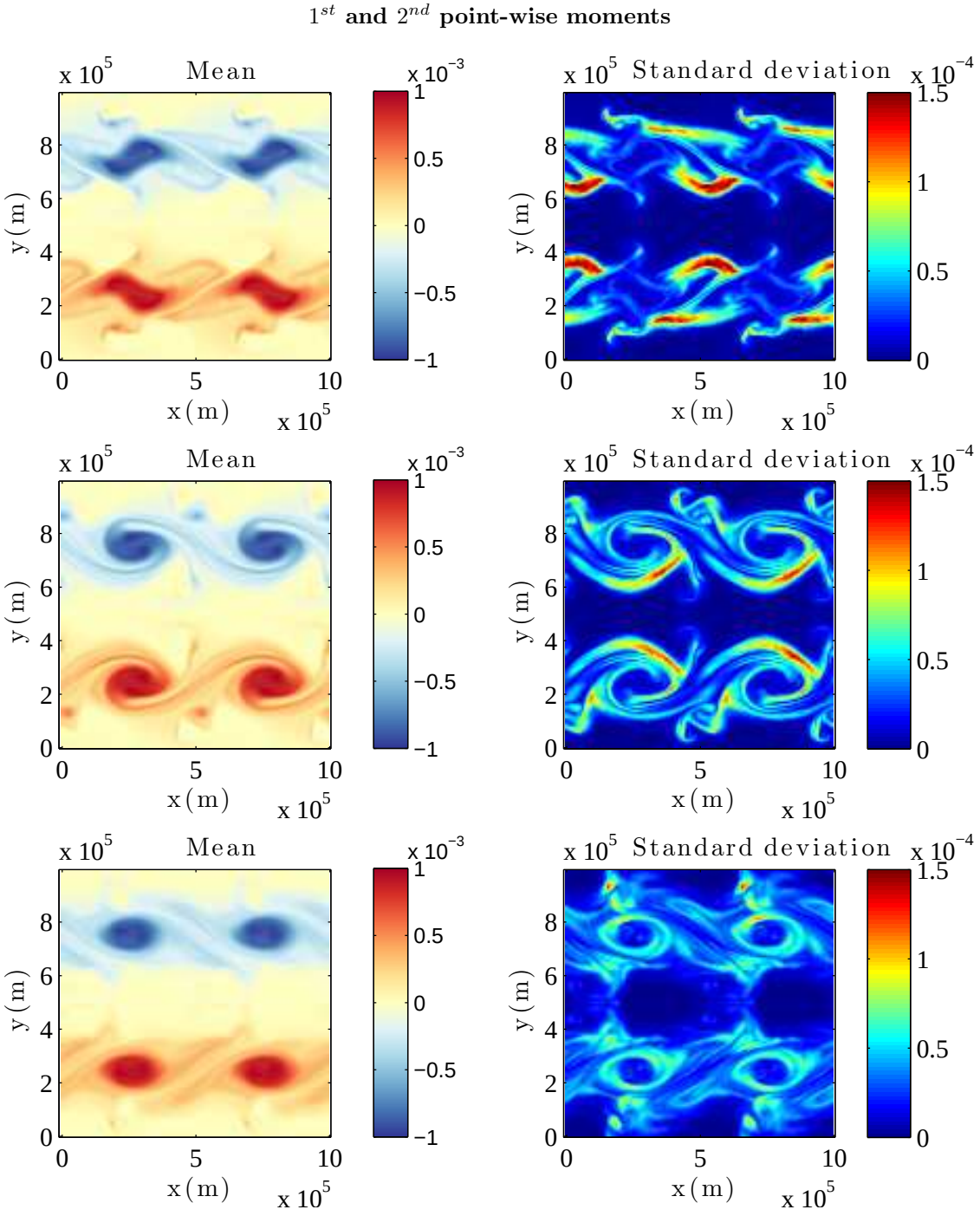

Figure 10: Point-wise mean (left) and standard deviation (right) of the buoyancy ($m.s^{-2}$) at $t = 17, 20$ and 30 days of advection for SQG_{MU} model at resolution 128^2. The moments are computed through MCMC simulations. The mean field is smoother than the individual realizations. Areas of higher variance appear first near the stretched saddle points. Then, at $t = 17$ days, the filament instabilities are triggered by the unresolved velocity component. The appearance of "pearl necklaces" can be observed. At $t = 20$, mushrooms-like structures also develop in the variance field near $(x, y) = (0, 100), (500, 100), (0, 900)$ and $(500, 900)$ (in km). At $t = 30$ days, these random structures are transported by the zonal jets.

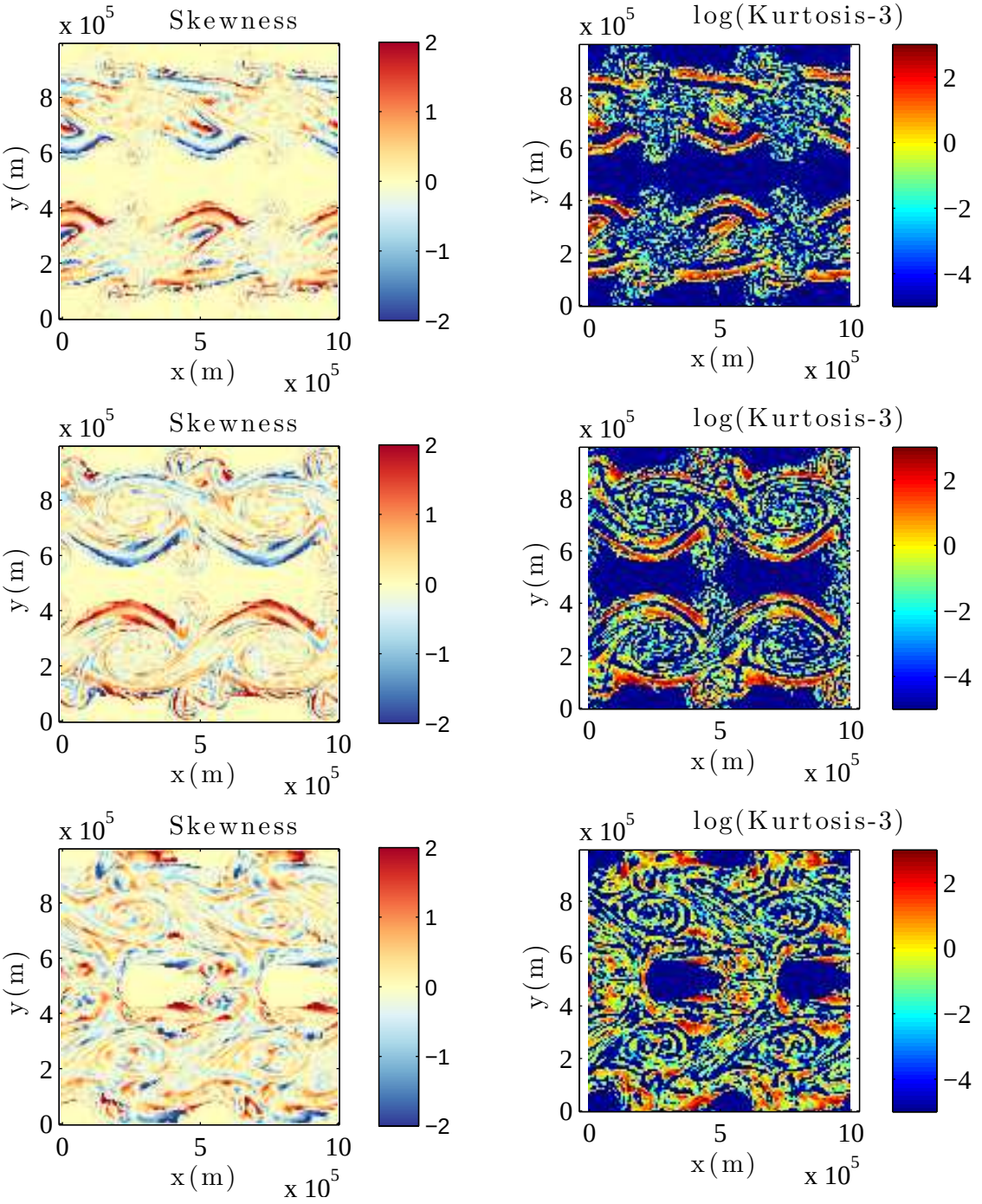

Figure 11: Point-wise skewness, and logarithm of the excess kurtosis of the buoyancy at $t = 17, 20$ and 30 days of advection for $\mathrm{SQG_{MU}}$ model at resolution 128^2. The moments are computed through MCMC simulations. Negative excess kurtosis is set to 0. The point-wise law of the tracer is clearly non-Gaussian. The "pearl necklace" events identified in the variance plots leads to fat-tailed distributions with skewness at $t = 17$ and 20 days. The random structures, associated with fat tails are then trapped in the zonal jets.

source/sink terms. Two of them are smooth in time and cancel out for homogeneous turbulence. The last forcing term – a random enstrophy source – is related to the angle between stable directions of resolved and unresolved velocities. Similarly to the deterministic case, a uniform PV yields a randomized SQG model, called SQG_{MU}, where the buoyancy is transported in the stochastic sense.

Simulation results are considered for the SQG_{MU} model which is a good representation of the transport under location uncertainty. As such, results are believed to hold for any fluid dynamics models under location uncertainty. As found, SQG_{MU} better resolves small-scale tracer structures than a usual SQG model simulated at the same resolution. The prescribed balance between noise and diffusion has also been confirmed. As further highlighted, an ensemble of simulations was able to estimate the amplitude and the position of its own errors in both spatial and spectral domains. This result suggests that the proposed randomized dynamics should be well suited for filtering and other data assimilation methods. On the contrary, a deterministic model with randomized initial conditions, either creates strong errors in its realizations (one order of magnitude larger than the unperturbed deterministic dynamics), or underestimates its own errors (one order of magnitude too low). A MATLAB® code simulating the SQG_{MU} model is available online (`http://vressegu.github.io/sqgmu`).

As a discussion, we can address the problem of uncertainty quantification (UQ) of an unresolved dynamics from an opposite point of view as the usual setting. Instead of specifying a form for the sub-grid velocity, we can wonder what is the optimal form of SPDE for UQ in fluid dynamics. As demonstrated, randomization of initial conditions is far from being sufficient to quantify uncertainty. Therefore, a random forcing is needed to inject randomness at each time step. The simplest choice is a forcing uncorrelated in time. Otherwise, additional stochastic equations need to be simulated to sample a time-correlated process. This is not desirable in high dimension and the correlation time of the process is often small anyway (Berner et al., 2011). A forcing uncorrelated in time is a source of energy. So, to be physically acceptable, the SPDE should involve a dissipative term to exactly compensate this source, even in non-stationary regime. The simplest choices of dissipation are diffusion and linear drag. For small-scale processes, the first is more suitable. Now, what is the form of a noise which brings as much energy as a diffusion removes? The proposed approach constitutes a suitable solution toward this goal.

To further improve the accuracy of the UQ, spatial inhomogeneity of the variance tensor $\boldsymbol{a}$ can be introduced from data or from additional models, as discussed in Resseguier et al. (2017a). This inhomogeneity may reduce possible spurious oscillations of tracer stable isolines. Such oscillations are visible on Figure 3 on the sides of the largest vortices. The assumption of time decorrelation may also be a limitation. Nevertheless, as shown by the numerical simulations, the method already achieve very good outcomes with an homogeneous noise component and no real time-scale separation between the resolved and unresolved velocities. Note in particular that since the noise is multiplicative, the random forcing is inhomogeneous even for homogeneous small-scale velocity.

Resseguier et al. (2017b) focuses on a system with a clear time-scale separation between the meso and sub-meso scale dynamics to explore the consequences of the QG assumptions under a strong uncertainty assumption ($\Upsilon \sim R_o$). A zero PV directly appears in the fluid interior and the horizontal velocity becomes divergent. This divergence provides a simple diagnosis of the frontolysis on warm sides of fronts and frontogenesis on cold sides of fronts.

Future works shall also focus on the potential benefits of the stochastic transport for data assimilation issues. As foreseen, the proposed stochastic formalism opens new horizons for ensemble

forecasting techniques and other UQ based dynamical approaches (e.g. Ubelmann et al., 2015). This stochastic setup has also been used to characterize chaotic transitions associated with breaking symmetries, also demonstrating interesting perspectives in that context.

Acknowledgments

The authors thank Aurélien Ponte, Jeroen Molemaker and Jonathan Gula for helpful discussions. We also acknowledge the support of the ESA DUE GlobCurrent project, the "Laboratoires d'Excellence" CominLabs, Lebesgue and Mer through the SEACS project.

References

J. Anderson and S. Anderson. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12):2741–2758, 1999.

J. Berner, G. Shutts, M. Leutbecher, and T. Palmer. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences*, 66(3):603–626, 2009.

J. Berner, S.-Y. Ha, J. Hacker, A. Fournier, and C. Snyder. Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Monthly Weather Review*, 139(6):1972–1995, 2011.

J. Berner, U. Achatz, L. Batte, A. De La Camara, D. Crommelin, H. Christensen, M. Colangeli, S. Dolaptchiev, C. Franzke, P. Friederichs, P. Imkeller, H. Jarvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajan, T. Palmer, C. Penland, J.-S. Von Storch, M. Sakradzija, M. Weniger, A. Weisheimer, P. Williams, and J.-I. Yano. Stochastic parameterization: towards a new view of weather and climate models. Technical report, arXiv:1510.08682 [physics.ao-ph], 2015.

W. Blumen. Uniform potential vorticity flow: part I. theory of wave interactions and two-dimensional turbulence. *Journal of the Atmospheric Sciences*, 35(5):774–783, 1978.

G. Boccaletti, R. Ferrari, and B. Fox-Kemper. Mixed layer instabilities and restratification. *Journal of Physical Oceanography*, 37(9):2228–2250, 2007.

M. Bocquet, P. Raanes, and A. Hannart. Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation. *Nonlin. Processes Geophys*, 22:645–662, 2015.

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

J. Chasnov. Simulation of the Kolmogorov inertial subrange using an improved subgrid model. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 3(1):188–200, 1991.

P. Constantin, A. Majda, and E. Tabak. Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar. *Nonlinearity*, 7(6):1495, 1994.

P. Constantin, Q. Nie, and N. Schörghofer. Front formation in an active scalar equation. *Physical Review E*, 60(3):2858, 1999.

P. Constantin, M. Lai, R. Sharma, Y. Tseng, and J. Wu. New numerical results for the surface quasi-geostrophic equation. *Journal of Scientific Computing*, 50(1):1–28, 2012.

G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992. ISBN 9780521385299.

F. Flandoli. The interaction between noise and transport mechanisms in PDEs. *Milan Journal of Mathematics*, 79(2):543–560, 2011.

C. Franzke, A. Majda, and E. Vanden-Eijnden. Low-order stochastic mode reduction for a realistic barotropic model climate. *Journal of the atmospheric sciences*, 62(6):1722–1745, 2005.

C. Franzke, T. O'Kane, J. Berner, P. Williams, and V. Lucarini. Stochastic climate theory and modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 6(1):63–78, 2015.

K. Gawędzki and A. Kupiainen. Anomalous scaling of the passive scalar. *Physical review letters*, 75(21):3834, 1995.

P. Gent and J. Mcwilliams. Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, 20(1):150–155, 1990.

H. Giordani, L. Prieur, and G. Caniaux. Advanced insights into sources of vertical velocity in the ocean. *Ocean Dynamics*, 56(5-6):513–524, 2006.

D. Givon, R. Kupferman, and A. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55, 2004.

G. Gottwald and J. Harlim. The role of additive and multiplicative noise in filtering complex dynamical systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 469(2155):20130096, 2013.

G. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 469(2156), 2013.

G. Gottwald, D. Crommelin, and C. Franzke. Stochastic climate theory. In *Nonlinear and Stochastic Climate Dynamics*. Cambridge University Press, 2015.

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

I. Held, R. Pierrehumbert, S. Garner, and K. Swanson. Surface quasi-geostrophic dynamics. *Journal of Fluid Mechanics*, 282:1–20, 1995.

S. Keating, S. Smith, and P. Kramer. Diagnosing lateral mixing in the upper ocean with virtual tracers: Spatial and temporal resolution dependence. *Journal of Physical Oceanography*, 41(8): 1512–1534, 2011.

P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1999.

R. Kraichnan. Small-scale structure of a scalar field convected by turbulence. *Physics of Fluids (1958-1988)*, 11(5):945–953, 1968.

R. Kraichnan. Anomalous scaling of a randomly advected passive scalar. *Physical review letters*, 72(7):1016, 1994.

H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

T. Kurtz. A limit theorem for perturbed operator semigroups with applications to random evolutions. *Journal of Functional Analysis*, 12(1):55–67, 1973.

G. Lapeyre and P. Klein. Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory. *Journal of physical oceanography*, 36(2):165–176, 2006.

C. Leith. Atmospheric predictability and two-dimensional turbulence. *Journal of the Atmospheric Sciences*, 28(2):145–161, 1971.

A. Majda and P. Kramer. Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena. *Physics report*, 314:237–574, 1999.

A. Majda, I. Timofeyev, and E. Eijnden. Models for stochastic climate prediction. *Proceedings of the National Academy of Sciences*, 96(26):14687–14691, 1999.

A. Majda, Ilya Timofeyev, and Eric Vanden Eijnden. A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics*, 54(8):891–974, 2001.

A. Majda, C. Franzke, and B. Khouider. An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366(1875):2427–2453, 2008.

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

R. Mikulevicius and B. Rozovskii. Stochastic Navier–Stokes equations for turbulent flows. *SIAM Journal on Mathematical Analysis*, 35(5):1250–1310, 2004.

L. Mitchell and G. Gottwald. Data assimilation in slow-fast systems using homogenized climate models. *Journal of the atmospheric sciences*, 69(4):1359–1377, 2012.

S. Orszag. Analytical theories of turbulence. *Journal of Fluid Mechanics*, 41(02):363–386, 1970.

G. Papanicolaou and W. Kohler. Asymptotic theory of mixing stochastic ordinary differential equations. *Communications on Pure and Applied Mathematics*, 27(5):641–668, 1974.

C. Penland and L. Matrosova. A balance condition for stochastic numerical models with application to the El Nino-southern oscillation. *Journal of climate*, 7(9):1352–1372, 1994.

C. Penland and P. Sardeshmukh. The optimal growth of tropical sea surface temperature anomalies. *Journal of climate*, 8(8):1999–2024, 1995.

R. Pierrehumbert and H. Yang. Global chaotic mixing on isentropic surfaces. *Journal of the atmospheric sciences*, 50(15):2462–2480, 1993.

C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*, volume 1905. Springer, 2007.

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part I: Random transport and general models. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017a.

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part III: SQG and frontal dynamics under strong turbulence. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017b.

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

C. Snyder, T. Bengtsson, and M. Morzfeld. Performance bounds for particle filters using the optimal proposal. *Monthly Weather Review*, 143:4750–4761, 2015.

C. Ubelmann, P. Klein, and L.-L. Fu. Dynamic interpolation of sea surface height and potential applications for future high-resolution altimetry mapping. *Journal of Atmospheric and Oceanic Technology*, 32(1):177–184, 2015.

G. Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press, 2006.

A Non-dimensional Boussinesq equations

To derive a non-dimensional version of the Boussinesq equations under location uncertainty (Resseguier et al., 2017a), we scale the horizontal coordinates $\tilde{\boldsymbol{x}}_h = L\boldsymbol{x}_h$, the vertical coordinate $\tilde{z} = hz$, the aspect ratio $D = h/L$ between the vertical and horizontal length scales. A characteristic time $\tilde{t} = Tt$ corresponds to the horizontal advection time U/L with horizontal velocity $\tilde{\boldsymbol{u}} = U\boldsymbol{u}$. A vertical velocity $\tilde{w} = (h/L)Uw$ is deduced from the divergence-free condition. We further take a scaled buoyancy $\tilde{b} = Bb$, pressure $\tilde{\phi}' = \Phi\phi'$ (with the density scaled pressures $\phi' = p'/\rho_b$ and $\mathrm{d}_t\phi_\sigma = \mathrm{d}_t p_\sigma/\rho_b$), and the earth rotation $\boldsymbol{f}^* = f\boldsymbol{k}$. For the uncertainty variables, we consider a horizontal uncertainty $\tilde{\boldsymbol{a}}_H = A_u\ \boldsymbol{a}_H$ corresponding to the horizontal 2×2 variance tensor; a vertical uncertainty vector $\tilde{a}_{zz} = A_w a_{zz}$ and a horizontal-vertical uncertainty vector $\tilde{\boldsymbol{a}}_{Hz} = \sqrt{A_u A_w}\boldsymbol{a}_{Hz}$ related to the variance between the vertical and horizontal velocity components. The resulting non-dimensional Boussinesq system under location uncertainty becomes:

Nondimensional Boussinesq equations under location uncertainty

Momentum equations

$$\mathrm{d}_t\boldsymbol{u} + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)\boldsymbol{u} + \left(\frac{R_o}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z\boldsymbol{u}$$
$$- \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}\boldsymbol{u}\right)\mathrm{d}t + O\left(\frac{R_o}{\Upsilon B_u}\right) + \frac{1}{R_o}\left(f_0 + R_o\beta y\right)\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)$$
$$= -E_u\,\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{38a}$$

$$\mathrm{d}_t w + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,w\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)w + \left(\frac{R_o}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z w$$
$$- \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}w\right)\mathrm{d}t + O\left(\frac{R_o}{\Upsilon B_u}\right) = \frac{\Gamma}{D^2}b\mathrm{d}t - \frac{E_u}{D^2}\,\partial_z\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{38b}$$

Buoyancy equation

$$\mathrm{d}_t b + \left(\boldsymbol{w}^*_\Upsilon\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)\right)\cdot\boldsymbol{\nabla}b - \frac{1}{2}\frac{1}{\Upsilon}\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{a}_H\boldsymbol{\nabla}b\right)\mathrm{d}t + O\left(\frac{R_o}{\Upsilon B_u}\right)$$
$$+ \frac{1}{(Fr)^2}\frac{1}{\Gamma}\left(w^*_{\Upsilon/2}\mathrm{d}t + \left(\frac{R_o}{B_u}\right)\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\right) = 0, \tag{38c}$$

Effective drift

$$\boldsymbol{w}^*_\Upsilon = \left(\boldsymbol{u}^*_\Upsilon, w^*_\Upsilon\right)^T,$$
$$= \left(\left(\boldsymbol{w} - \frac{1}{2\Upsilon}\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right), \left(w - \left(\frac{R_o}{2\Upsilon B_u}\right)\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_{Hz} + O\left(\frac{R_o}{\Upsilon B_u}\right)^2\right)\right)^T, \tag{38d}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0, \tag{38e}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = 0, \tag{38f}$$

$$\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H\right)^T + 2\frac{R_o}{B_u}\boldsymbol{\nabla}_H\cdot\partial_z\boldsymbol{a}_{Hz} + O\left(\left(\frac{R_o}{B_u}\right)^2\right) = 0. \tag{38g}$$

Here, we do not separate the time-correlated components and the time-uncorrelated components in the momentum equations. The terms in $O\left(\frac{R_o}{B_u}\right)$ and $O\left(\frac{R_o}{B_u}\right)^2$ are related to the time-uncorrelated vertical velocity. These terms are too small to appear in the final QG model ($B_u = O(1)$ in QG approximation) and not explicitly shown. We only make appear the big O approximations. Traditional non-dimensional numbers are introduced : the Rossby number $R_o = U/(f_0 L)$ with f_0 the average Coriolis frequency; the Froude number ($F_r = U/(Nh)$), ratio between the advective time to the buoyancy time; E_u, the Euler number, ratio between the pressure force and the inertial forces, $\Gamma = Bh/U^2 = D^2BT/W$ the ratio between the mean potential

energy to the mean kinetic energy. To scale the buoyancy equation, the ratio between the buoyancy advection and the stratification term has also been introduced:

$$\frac{B/T}{N^2W} = \frac{B}{N^2h} = \frac{U^2}{N^2h^2}\frac{Bh}{U^2} = Fr^2\Gamma. \tag{39}$$

Besides those traditional dimensionless numbers, this system introduces Υ, relating the large-scale kinetic energy to the energy dissipated by the unresolved component:

$$\Upsilon = \frac{UL}{A_u} = \frac{U^2}{A_u/T}. \tag{40}$$

B QG model under moderate uncertainty

Hereafter, we consider the QG approximation ($Ro \ll 1$ and $Bu \sim 1$), for $\Upsilon \sim 1$. We focus on solutions of the Boussinesq model with Rossby number going to zero. To derive the evolution equations corresponding to this limit, the solution of the non-dimentional Boussinesq model (Appendix A) is developed as a power series of the Rossby number:

$$\begin{pmatrix} \boldsymbol{w} \\ b \\ \phi \end{pmatrix} = \sum_{k=0}^{\infty} {Ro}^k \begin{pmatrix} \boldsymbol{w}_k \\ b_k \\ \phi_k \end{pmatrix}. \tag{41}$$

According to the horizontal momentum equation (38a), the scaling of the pressure still corresponds to the usual geostrophic balance. This sets the Euler number as:

$$Eu \sim \frac{1}{Ro}. \tag{42}$$

For the ocean, the aspect ratio, $D \triangleq H/L$, is small and $D^2 \ll 1$. As a consequence,

$$\frac{D^2}{Eu} \sim D^2 Ro \ll D^2 \ll 1 \text{ and } \frac{D^2}{Eu\Upsilon} \sim \frac{D^2 Ro}{\Upsilon} \leqslant D^2 \ll 1. \tag{43}$$

Therefore, the inertial and diffusion terms are negligible in the vertical momentum equation. The hydrostatic assumption is still valid. This leads to the classical QG scaling of the buoyancy equation:

$$\Gamma \sim Eu \sim \frac{1}{Ro} \text{ and } \frac{1}{{Fr}^2\Gamma} \sim \frac{Ro}{{Fr}^2} = \frac{Bu}{Ro}. \tag{44}$$

In the following, the subscript H is omitted for the differential operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ. They all represent 2D operators. Only keeping terms of order 0 and 1, we get the following system:

Momentum equations

$$Ro\left(\mathrm{d}_t\boldsymbol{u}+(\boldsymbol{u}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\boldsymbol{u}-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}\boldsymbol{u}\right)\mathrm{d}t+O\left(\frac{Ro}{\Upsilon Bu}\right)\right)$$
$$+\left(f_0+Ro\beta y\right)\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)=-\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{45}$$

$$b\,\mathrm{d}t+O\left(RoD^2\right)=\partial_z\left(\phi'\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{46}$$

Buoyancy equation

$$\frac{Ro}{Bu}\left(\mathrm{d}_tb+\boldsymbol{\nabla}b\cdot\left(\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)_H\right)+\partial_zb\;w\mathrm{d}t-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}b\right)\mathrm{d}t\right)$$
$$+w\mathrm{d}t-\frac{1}{\Upsilon}\frac{Ro}{Bu}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t+\frac{Ro}{Bu}\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z+O\left(\frac{{Ro}^2}{\Upsilon{Bu}^2}\right)=0, \tag{47}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}+\partial_zw=0, \tag{48}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)_H+\frac{Ro}{Bu}\partial_z\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)_z=0, \tag{49}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T+2\frac{Ro}{Bu}\boldsymbol{\nabla}\cdot\partial_z\boldsymbol{a}_{Hz}+O\left(\left(\frac{Ro}{Bu}\right)^2\right)=0. \tag{50}$$

The thermodynamic equation (47) at 0 order leads to :

$$w_0=0, \tag{51}$$

and then, by the large-scale incompressibility equation (48), the 0-order horizontal velocity is divergence-free. Following the scaling assumption, the horizontal small-scale velocity is also divergence-free (49). The horizontal momentum equation (45) at the 0-th order leads to:

$$\boldsymbol{u}_0=\frac{1}{f_0}\,\boldsymbol{\nabla}^\perp\phi_0'\text{ and }(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=\frac{1}{f_0}\,\boldsymbol{\nabla}^\perp\mathrm{d}_t\phi_\sigma, \tag{52}$$

where time-correlated and time-uncorrelated components have been separated by the mean of uniqueness of the semi-martingale decomposition (Kunita, 1997). Being divergent-free, both components can be expressed with two stream functions ψ_0 and $\mathrm{d}_t\psi_\sigma$:

$$\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp\psi_0\text{ and }(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=\boldsymbol{\nabla}^\perp\mathrm{d}_t\psi_\sigma, \tag{53}$$

exactly corresponding to the dimensionless pressure terms:

$$\psi_0=\frac{1}{f_0}\,\phi_0'\text{ and }\mathrm{d}_t\psi_\sigma=\frac{1}{f_0}\,\mathrm{d}_t\phi_\sigma. \tag{54}$$

Deriving these equations along z and introducing the hydrostatic equilibrium (46) – decomposed between correlated and uncorrelated components – yields the classical thermal wind balance at

large-scale for the 0-th order terms. The buoyancy variable does not involve any white noise term, and the small-scale random velocity is thus almost constant along z, as

$$\partial_z \boldsymbol{u}_0 = \frac{1}{f}_0 \boldsymbol{\nabla}^\perp b_0 \text{ and } \partial_z (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H = O\left(R_o D^2\right). \tag{55}$$

Accordingly the variance tensor scales as:

$$\forall i,j \in H,\ \partial_z a_{ij} = O\left({R_o}^2 D^4\right), \tag{56}$$

which is negligible in all equations, and the uncertain random field solely depends on the horizontal coordinates. Since $R_o/B_u \sim R_o$, the 1-st order term of the buoyancy equation must be kept to describe the evolution of b_0:

$$\frac{1}{B_u}\mathbb{D}_{0t}^H b_0 + w_1 \mathrm{d}t - \frac{1}{\Upsilon}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T \mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z = 0, \tag{57}$$

where, for all functions h,

$$\mathbb{D}_{0t}^H h \triangleq \mathrm{d}_t h + \boldsymbol{\nabla} h \cdot \left(\boldsymbol{u}_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\right) - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial_{ij}^2\left(a_{ij}h\right)\mathrm{d}t. \tag{58}$$

Taking the derivative along z leads to:

$$\begin{aligned}\frac{1}{B_u}\left(\mathbb{D}_{0t}^H \partial_z b_0 + \boldsymbol{\nabla} b_0 \cdot\ \partial_z\left(\boldsymbol{u}_0\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma} d\boldsymbol{B}_t)_H\right) - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial_{ij}^2\left(\partial_z a_{ij} b_0\right)\mathrm{d}t\right)& \\ + \partial_z w_1 \mathrm{d}t - \frac{1}{\Upsilon}\left(\boldsymbol{\nabla}\cdot\partial_z a_{Hz}\right)^T \mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\partial_z(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z = 0.&\end{aligned} \tag{59}$$

The introduction of the thermal wind equations (55) and incompressibility conditions (48-50) helps simplifying this equation as:

$$\frac{1}{B_u}\mathbb{D}_{0t}^H \partial_z b_0 \ - \ \boldsymbol{\nabla}\cdot\boldsymbol{u}_1 \mathrm{d}t \ + \ \left(\frac{R_o}{B_u}\right)^{-1}\frac{1}{\Upsilon}\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T \mathrm{d}t \ - \ \left(\frac{R_o}{B_u}\right)^{-1}\frac{1}{\Upsilon^{1/2}}\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H \ = \ 0. \tag{60}$$

Note the factor $\left(\frac{R_o}{B_u}\right)^{-1}$ appears. It comes from the incompressible conditions (49) and (50), leading $\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$ and $\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T \mathrm{d}t$ to both scale as $\frac{R_o}{B_u}$. The hydrostatic balance at 0-order links the buoyancy to the pressure, and then to the stream function

$$\partial_z b_0 = \partial_z^2 \phi_0 = f_0 \partial_z^2 \psi_0. \tag{61}$$

The 1-st order term of the vertical velocity is not known. Yet, the system can be closed using the vorticity equation at order 1:

$$\boldsymbol{\nabla}^\perp\cdot\left(\mathbb{D}_{0t}^H \boldsymbol{u}_0\right) \ + \ f_0\left(\boldsymbol{\nabla}\cdot\boldsymbol{u}_1 + \left(\frac{R_o}{B_u}\right)^{-1}\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\right) + \boldsymbol{\nabla}(\beta y)\cdot\left(\boldsymbol{u}_0\mathrm{d}t + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\right) \ = \ 0, \tag{62}$$

where the divergence terms come from the constant Coriolis term.

Again, factors $\left(\frac{R_o}{B_u}\right)^{-1}$ compensate the order of magnitude of $\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$ and $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H)^T\,\mathrm{d}t$. Then,

$$\mathbb{D}_{0t}^H(\Delta\psi_0)+f_0\boldsymbol{\nabla}\cdot u_1\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\left(\frac{R_o}{B_u}\right)^{-1}f_0\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H+\beta\left(v_0\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_y\right)$$
$$+\frac{1}{\Upsilon^{1/2}}\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}_0^T\right)-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}_0\right)\mathrm{d}t=0. \quad (63)$$

Using (60) and (61), we get:

$$\mathbb{D}_{0t}^H\left(\Delta\psi_0+f_0+\beta y+\frac{f_0^2}{B_u}\partial_z^2\psi_0\right)=-\boldsymbol{\nabla}\cdot a_{Hy}\beta\mathrm{d}t-f_0\left(\frac{R_o}{B_u}\right)^{-1}\frac{1}{2\Upsilon}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H)^T\,\mathrm{d}t$$
$$-\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}_0^T\right)+\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}_0\right)\mathrm{d}t. \quad (64)$$

We recall that coefficients $\left(\frac{R_o}{B_u}\right)^{-1}$ are still present since

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\sim\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H)^T\,\mathrm{d}t\sim\left(\frac{R_o}{B_u}\right). \quad (65)$$

If we rewrite the equation with dimensional quantities, the evolution equation for $u_0=\lim_{R_o\to0}u$ is obtained (dropping the index 0 for clarity):

$$\mathbb{D}_t^HQ\;=\;-\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right)+\frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\mathrm{d}t-\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_Hf))^T\,\mathrm{d}t, \quad (66)$$

where Q is the QG potential vorticity:

$$Q\triangleq\Delta\psi+f+\left(\frac{f_0}{N}\right)^2\partial_z^2\psi. \quad (67)$$

Note, (53) provides the geostrophic balance for the small-scale velocity component. To express the material derivative of Q, the noise term is expanded:

$$-\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right)=-\sum_{k,j\in H}\partial_{kj}^2\psi\partial_k\boldsymbol{\sigma}_{j\bullet}\mathrm{d}\boldsymbol{B}_t. \quad (68)$$

According to Resseguier et al. (2017a), the difference between the material derivative, D_tQ, and the stochastic transport operator $\mathbb{D}_tQ$, is a function of the time-uncorrelated forcing:

$$\left\{\begin{array}{rcl}\mathbb{D}_tQ&=&f_1\mathrm{d}t+\boldsymbol{h}_1^T\mathrm{d}\boldsymbol{B}_t,\\D_tQ&=&f_2\mathrm{d}t+\boldsymbol{h}_2^T\mathrm{d}\boldsymbol{B}_t,\end{array}\right.\iff\left\{\begin{array}{rcl}f_2&=&f_1+\mathrm{tr}\big((\boldsymbol{\sigma}^T\boldsymbol{\nabla})\,\boldsymbol{h}_1^T\big),\\\boldsymbol{h}_1&=&\boldsymbol{h}_2.\end{array}\right. \quad (69)$$

The expression of h_1 is given by equation (66) and the above formulas give:

$$D_tQ - \mathbb{D}_tQ = \sum_{i\in H} \boldsymbol{\sigma}_{i\bullet}\partial_i \left(- \sum_{j,k\in H} \partial_k \boldsymbol{\sigma}_{j\bullet}\partial^2_{kj}\psi \right)^T, \tag{70}$$

$$= - \sum_{i,j,k\in H} \left(\boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{kj}\psi + \boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^3_{ijk}\psi \right). \tag{71}$$

With the use of the small-scale incompressibility, we obtain:

$$\frac{1}{2}\sum_{i,j\in H} \partial^2_{ij}\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right) =$$

$$\sum_{i,j,k\in H} \left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{ik}\psi + \boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{jk}\psi + \boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^3_{ijk}\psi \right). \tag{72}$$

From (71) and (72), it yields:

$$D_tQ - \left(\mathbb{D}_tQ - \frac{1}{2}\sum_{i,j\in H} \partial^2_{ij}\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\right) = \sum_{i,j,k\in H} \left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{ik}\psi \right). \tag{73}$$

Denoting, α, the following matrix

$$\alpha_{ij} \stackrel{\triangle}{=} \sum_{k\in H} \partial_k\boldsymbol{\sigma}_{i\bullet}\partial_j\boldsymbol{\sigma}^T_{k\bullet} = \sum_{k\in H} \partial_k(\boldsymbol{\sigma}_{i\bullet}\partial_j\boldsymbol{\sigma}^T_{k\bullet}), \tag{74}$$

we have

$$\boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi) = \sum_{i,j,k\in H} \left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{ik}\psi \right), \tag{75}$$

$$= D_tQ - \left(\mathbb{D}_tQ - \frac{1}{2}\sum_{i,j\in H} \partial^2_{ij}\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\right), \tag{76}$$

and the material derivative of the PV finally reads:

$$D^H_t Q = \boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_H f)\right)^T \mathrm{d}t - \mathrm{tr}\left[\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T_H \boldsymbol{\nabla}\boldsymbol{u}^T\right]. \tag{77}$$

To note, the transpose of the matrix α has a compact expression:

$$\alpha^T = \sum_p \left(\boldsymbol{\nabla}\boldsymbol{\sigma}^T_{Hp}\right)^2. \tag{78}$$

To better assess the role of the random source term (the last term of (77)), it is decomposed in terms of symmetric and anti-symmetric parts of the small-scale/large-scale deformation tensors. Let us denote $\boldsymbol{S}$ and $\boldsymbol{S}_{\sigma\mathrm{d}B_t}$ the symmetric parts of $\boldsymbol{\nabla}\boldsymbol{u}^T$ and $\boldsymbol{\nabla}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T_H$, respectively. Associated with divergence-free velocities, these symmetric parts, so-called strain rate tensors, have zero trace.

Terms $-\frac{1}{2}\omega\boldsymbol{J}$ and $-\frac{1}{2}\omega_{\sigma \mathrm{d}B_t}\boldsymbol{J}$ will stand for the anti-symmetric parts, where $\boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the $\frac{\pi}{2}$ rotation. The factors ω and $\omega_{\sigma \mathrm{d}B_t}$ are the large-scale and the small-scale components of the vorticity, respectively. Using $\boldsymbol{JJ} = -\mathbb{I}_d$ and $\mathrm{tr}[\boldsymbol{MN}] = tr[\boldsymbol{NM}]$ yields:

$$
\begin{aligned}
-\mathrm{tr}\left[\boldsymbol{\nabla}^{\perp}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right] &= -\mathrm{tr}\left[\boldsymbol{J}\left(\boldsymbol{S}_{\sigma \mathrm{d}B_t} - \frac{1}{2}\omega_{\sigma \mathrm{d}B_t}\boldsymbol{J}\right)\left(\boldsymbol{S} - \frac{1}{2}\omega\boldsymbol{J}\right)\right], && (79)\\
&= -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right] - \frac{1}{2}\omega_{\sigma \mathrm{d}B_t}\underbrace{\mathrm{tr}\left[\boldsymbol{S}\right]}_{=0} \\
&\quad -\frac{1}{2}\,\omega\,\underbrace{\mathrm{tr}\left[\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right]}_{=0} + \frac{1}{4}\,\omega\,\omega_{\sigma \mathrm{d}B_t}\underbrace{\mathrm{tr}\left[\boldsymbol{J}\right]}_{=0}, && (80)\\
&= -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right]. && (81)
\end{aligned}
$$

This term thus only depends on the strain rate tensors of $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$. The PV transport can thus be rewritten as:

$$
D_t^H Q = \boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}_H f\right)\right)^T \mathrm{d}t - \mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right]. \tag{82}
$$

The noise term can be further expressed using the stable directions of the flows defined by $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$, respectively. In the following, we will omit writing the $\mathrm{d}\boldsymbol{B}_t$ factor. The two strain rate tensors are decomposed in orthogonal basis:

$$
\boldsymbol{S} = \boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^T = \sum_{p=1}^{2}\Xi_{pp}\boldsymbol{V}_{\bullet p}\boldsymbol{V}_{\bullet p}^T \text{ and } \boldsymbol{S}_{\sigma \mathrm{d}B_t} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^T, \tag{83}
$$

where $\boldsymbol{V}_{\bullet p}^T\boldsymbol{V}_{\bullet q} = \boldsymbol{W}_{\bullet p}^T\boldsymbol{W}_{\bullet q} = \delta_{pq}$, $\Xi_{11} = -\Xi_{22} < 0$ and $\Lambda_{11} = -\Lambda_{22} < 0$.

$$
\begin{aligned}
-\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right] &= -\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\mathrm{tr}\left[\boldsymbol{V}_{\bullet p}\boldsymbol{V}_{\bullet p}^T\boldsymbol{J}\boldsymbol{W}_{\bullet q}\boldsymbol{W}_{\bullet q}^T\right], && (84)\\
&= -\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\left(\boldsymbol{V}_{\bullet p}^T\boldsymbol{W}_{\bullet q}\right)\left(\boldsymbol{V}_{\bullet p}^T\boldsymbol{J}\boldsymbol{W}_{\bullet q}\right), && (85)\\
&= -\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\ \cos(\theta_{pq})\cos\left(\theta_{pq}+\frac{\pi}{2}\right), && (86)\\
&= \frac{1}{2}\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\ \sin(2\theta_{pq}), && (87)
\end{aligned}
$$

where $\theta_{pq} \overset{\triangle}{=} \widehat{(\boldsymbol{V}_{\bullet p}, \boldsymbol{W}_{\bullet q})}$ is the angle between $\boldsymbol{V}_{\bullet p}$ and $\boldsymbol{W}_{\bullet q}$. Using the relations between the eigenvalues and the orthogonality of the eigenvectors, it finally comes:

$$
\begin{aligned}
-\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right] &= \frac{1}{2}\Xi_{11}\Lambda_{11}\left(\sin\left(2\theta_{11}\right) - \sin\left(2\left(\theta_{11}-\frac{\pi}{2}\right)\right) - \sin\left(2\left(\theta_{11}+\frac{\pi}{2}\right)\right) + \sin\left(2\theta_{11}\right)\right), \\
&= 2\underbrace{\Xi_{11}\Lambda_{11}}_{>0}\sin\left(2\theta_{11}\right). && (88)
\end{aligned}
$$