



Theoretical and Computational Analysis of the Thermal Quasi-Geostrophic Model

D. Crisan¹ · D. D. Holm¹ · E. Luesink² · P. R. Mensah¹ · W. Pan¹

Received: 6 October 2022 / Accepted: 7 July 2023 / Published online: 16 August 2023
© The Author(s) 2023

Abstract

This work involves theoretical and numerical analysis of the thermal quasi-geostrophic (TQG) model of submesoscale geophysical fluid dynamics (GFD). Physically, the TQG model involves thermal geostrophic balance, in which the Rossby number, the Froude number and the stratification parameter are all of the same asymptotic order. The main analytical contribution of this paper is to construct local-in-time unique strong solutions for the TQG model. For this, we show that solutions of its regularised version α -TQG converge to solutions of TQG as its smoothing parameter $\alpha \rightarrow 0$ and we obtain blow-up criteria for the α -TQG model. The main contribution of the computational analysis is to verify the rate of convergence of α -TQG solutions to TQG solutions as $\alpha \rightarrow 0$, for example, simulations in appropriate GFD regimes.

Keywords Oceanography · Bathymetry · Potential vorticity · Analysis of partial differential equations · Finite element methods · Rossby waves

Mathematics Subject Classification 35R99 · 65D30 · 65M60 · 76U50 · 76U65 · 86A05

Communicated by Paul Newton.

✉ E. Luesink
e.luesink@utwente.nl

D. Crisan
d.crisan@imperial.ac.uk

D. D. Holm
d.holm@imperial.ac.uk

P. R. Mensah
p.mensah@imperial.ac.uk

W. Pan
wei.pan@imperial.ac.uk

¹ Department of Mathematics, Imperial College, London SW7 2AZ, UK

² Department of Mathematics, University of Twente, 7500 AE Enschede, The Netherlands

Abbreviations

TQG	Thermal quasi-geostrophic
GFD	Geophysical fluid dynamics
SSH	Sea-surface height
IL	Inhomogeneous layer
TRSW	Thermal rotating shallow water
L1	Lagrangian 1
SALT	Stochastic advection by Lie transport
PDE	Partial differential equation
$i\alpha$	Intermediate α
FEM	Finite element method
CG	Continuous Galerkin
DG	Discontinuous Galerkin
SSPRK3	Strong stability preserving Runge–Kutta 3
BKM	Beale–Kato–Majda

1 Introduction

1.1 Purpose

The thermal quasi-geostrophic (TQG) equations comprise a mesoscale model of ocean dynamics in a solution regime near thermal geostrophic balance. In thermal geostrophic balance, three forces—the Coriolis, hydrostatic pressure gradient and buoyancy-gradient forces—sum to zero. Numerical simulations of the TQG equations show the onset of instability at high wavenumbers which creates small coherent structures which resemble submesoscale (1–20 km) features observed in satellite ocean colour images as seen in Fig. 1.

On the left panel of Fig. 1, one sees submesoscale features which are prominently displayed in computational simulations of TQG equations for sea-surface height (SSH). The right panel of Fig. 1 shows the surface of the Lofoten Basin off the coast of Norway near the Faroe Islands. In crossing the Lofoten Basin, warm saline Atlantic waters create buoyancy fronts as they meet the cold currents of the Arctic Ocean. Figure 1 displays several features of submesoscale currents surveyed in McWilliams (2019). High-resolution (4 km) computational simulations of the Lofoten Vortex have recently discovered that its time-mean circulation is primarily barotropic, Volkov et al. (2015), thereby making the flow in the Lofoten Basin a reasonable candidate for investigation using vertically averaged dynamics such as the TQG dynamical system. Both images show a plethora of multiscale features involving shear interactions of vortices, fronts, plumes, spirals, jets and Kelvin–Helmholtz roll-ups. The submesoscale features persist and interact strongly with each other in Kelvin–Helmholtz roll-up dynamics, instead of simply cascading energy to higher wavenumbers. This observation means that the instabilities which create these submesoscale features quickly regain stability

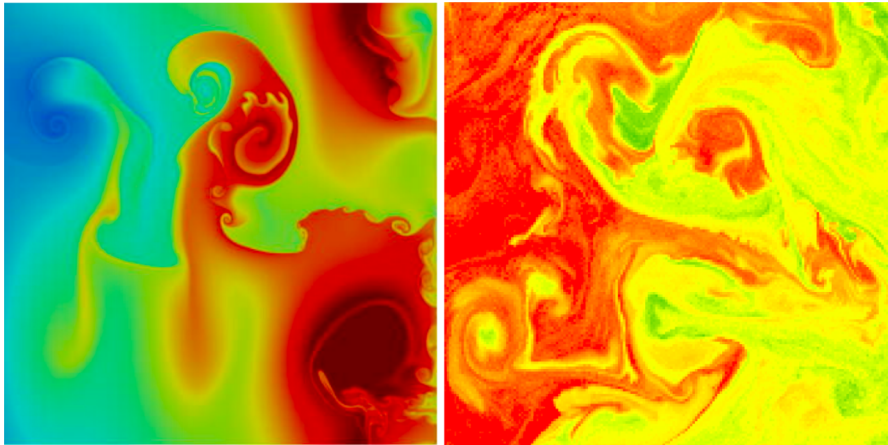


Fig. 1 Comparison of a computational simulation of TQG solutions with a satellite observation of ocean colour on a section of the Lofoten Vortex, courtesy of <https://ovl.oceandatalab.com/> which illustrates the configurations of submesoscale currents obtained from ESA Sentinel-3 OLCI instrument observations of chlorophyll on the surface of the Norwegian Sea in the Lofoten Basin, near the Faroe Islands

without cascading them to ever smaller scales. The present work aims to understand these features of the TQG solution dynamics, both analytically and numerically.

The contributions of this paper The main analytical contribution of this paper is the construction of a unique local strong solution of the TQG model. In particular, we show that there is a unique maximal solution of the TQG equation defined on a (possibly infinite) time interval $[0, T_{\max})$. The solution will be shown to exist in a suitable Sobolev space. More precisely, provided that the initial data resides in a chosen Sobolev space, the solution at time $t > 0$ will remain in this space as long as $t < T_{\max}$. Should $t < T_{\max}$ be finite, then the solution will blow up in the chosen Sobolev norm.

As a second analytical contribution, we show that the TQG model depends continuously on the initial condition. This dependence only holds in a slightly weaker norm than the one corresponding to the space where the solution resides. This is a useful property from a numerical perspective. It implies that initial small errors when simulating the TQG model will stay small at any subsequent time.

The third analytical contribution is to construct a regularised version of the TQG model, termed the α -TQG model. This model is constructed in a similar manner as the α model for the Euler and Navier–Stokes equations, see (Foias et al. 2001, 2002; Marsden and Shkoller 2003). The α model approach is a dispersive regularisation, rather than being dissipative. This feature has proven remarkably successful in applications of global ocean circulation, largely because its non-dissipative regularisation preserves the onset of baroclinic instability as shown in Holm and Wingate (2005). In the present case, the α -TQG equations are found to possess a unique maximal solution which is also only continuous with respect to the initial conditions in a larger space with weaker norm. In addition, we show that the α -TQG solution converges to the TQG solution as $\alpha \rightarrow 0$ in a norm that depends on two physical parameters, the vorticity and the

gradient of buoyancy. We also identify the rate of convergence as a function of the α -parameter on a time interval where both the TQG solution and the α -TQG solution are shown to exist for any $\alpha > 0$.

The blow-up phenomenon is important, particularly if it is observed in numerical simulations. Therefore, blow-up criteria (in other words, criteria required for the solution to blow up) are important. In this paper we state three characterisations of the blow-up time. The fourth analytical contribution of this paper is to show that blow-up occurs in the α -TQG model, if either:

- (1) the $L^\infty(\mathbb{T}^2)$ -norm of the buoyancy gradient ∇b blows up at $T_{\max} < \infty$;
- (2) the $L^\infty(\mathbb{T}^2)$ -norm of the velocity gradient $\nabla \mathbf{u}$ blows up at $T_{\max} < \infty$ or that;
- (3) the $W^{1,2}(\mathbb{T}^2)$ -norm of the buoyancy gradient ∇b blows up at $T_{\max} < \infty$.

The contribution of this paper from a numerical perspective is to describe our spatial and temporal discretisation methods for approximating TQG and α -TQG solutions, and to analyse aspects of numerical conservation properties with respect to the theoretical conserved quantities of the α -TQG system. Example simulation results are included and are used to verify numerically the theoretical convergence results. Additionally, we provide linear stability analysis results for the α -TQG system. Given that we have convergence of α -TQG solutions to TQG, the linear stability results can be viewed as generalisations of those shown in Holm et al. (2021) for the TQG system.

1.2 Brief History of the TQG Model

The history of the TQG model goes back about half a century, perhaps first elucidated by O'Brien and Reid (1967) as sketched in Beron-Vera (2021a). Briefly put, the TQG model generalises the classical QG equations by introducing horizontal gradients of buoyancy which alter the geostrophic balance to include the inhomogeneous thermal effects which influence buoyancy. Indeed, O'Brien and Reid (1967) write that their work was inspired by observations that the passage of hurricanes could draw enough heat from the ocean to significantly lower the sea-surface temperature in the Gulf of Mexico. Since O'Brien and Reid (1967) introduced their two-layer model, further developments of it have been applied to a variety of ocean processes, particularly to equatorial dynamics. For more details of the theoretical model developments, see (Ripa 1993, 1995, 1999) and for developments of applications in oceanography see (Anderson and McCreary Jr 1985; Beier 1997; McCreary Jr et al. 1997; Schopf and Cane 1983), as well as other citations in Beron-Vera (2021a). In particular, Ripa refers to the TQ models as inhomogeneous-layer (IL) models and his papers explain rational derivations of theories with increasing vertical structure IL1, IL2, etc. The TQG model analysed here and derived systematically from asymptotic expansions in small dimensionless parameters of the Hamilton's principle for the rotating, stratified Euler equations in Holm et al. (2021) is equivalent to the model IL0QG of Ripa (1996) recently analysed in Beron-Vera (2021b).

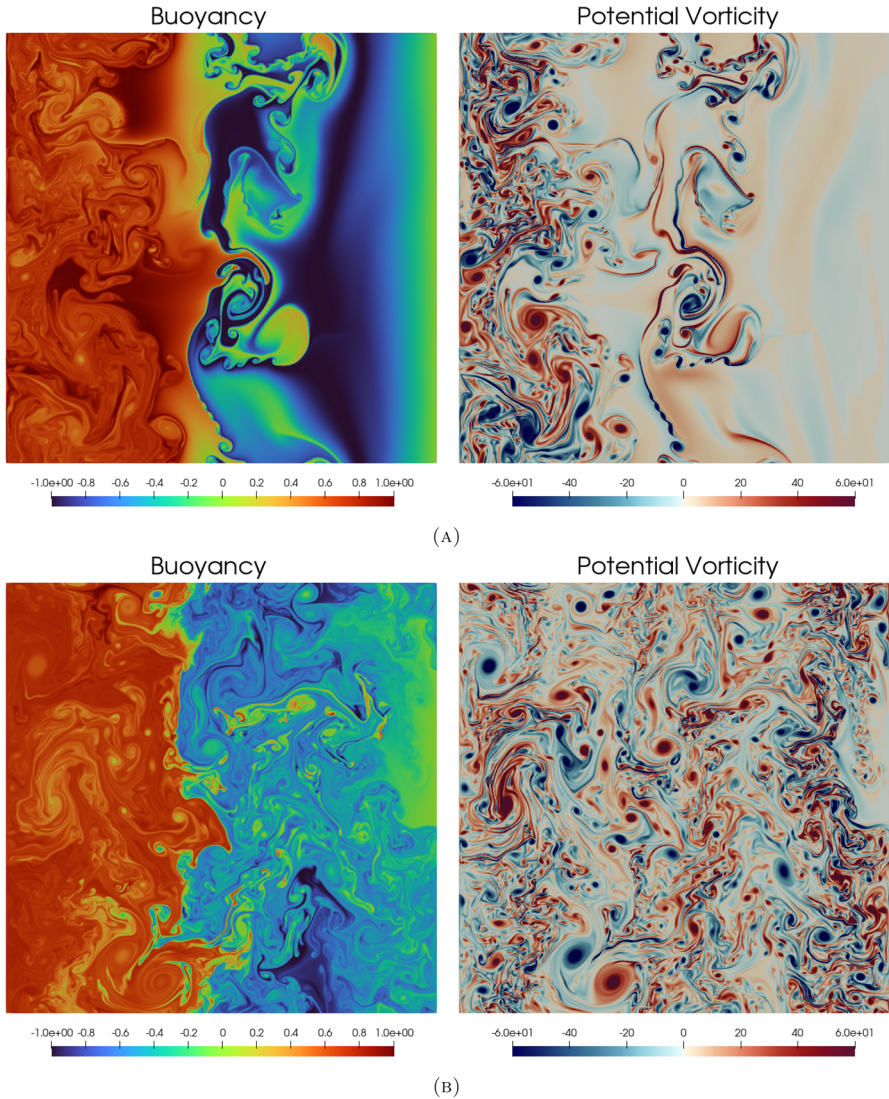


Fig. 2 Point-in-time snapshots of a computational simulation of the TQG system in a channel domain. The bathymetry function was chosen to be $h(x, y) = \cos(2\pi x) + \frac{1}{2} \cos(4\pi x) + \frac{1}{3} \cos(6\pi x)$. The initial potential vorticity was chosen to be an inverted Gaussian kernel $\omega(0, x, y) = -\exp(-\frac{1}{2}(2\pi y - \pi)^2)$. The initial buoyancy was chosen to be $b(0, x, y) = \sin(2\pi x)$. Subfigure 2a illustrates how the interaction between the gradients of buoyancy and bathymetry impacts the development of features during the early phases of the flow. Subfigure 2b illustrates the fully developed flow. Note that in the figure for potential vorticity in Subfigure 2b, the distribution of features at different scales is controlled by the chosen bathymetry function

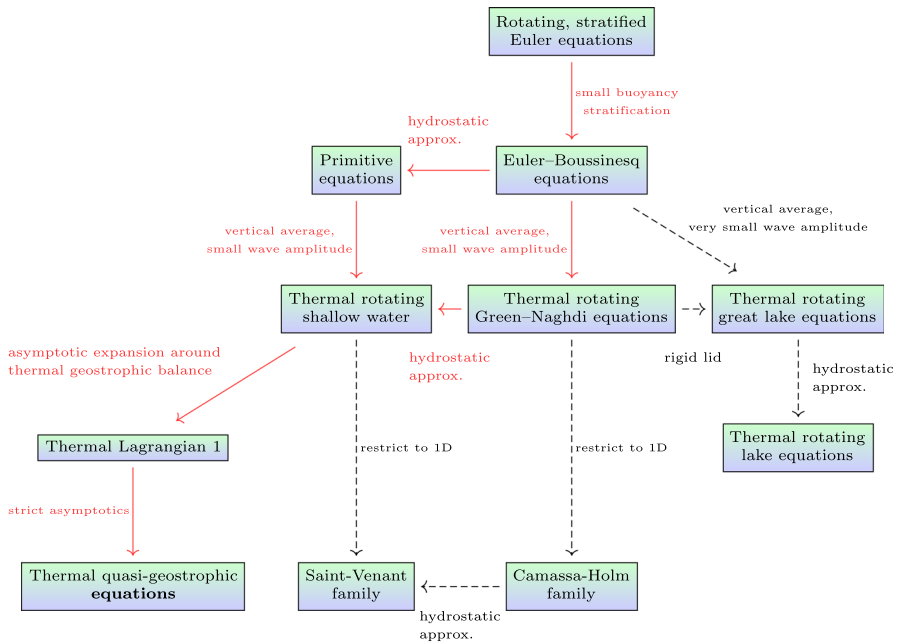


Fig. 3 The tree of model derivations can function as a roadmap of geophysical fluid dynamics. The solid red arrows indicate the sequence of approximations that lead to the thermal quasi-geostrophic equations

1.3 A Sketch of the Derivation of the TQG Model

We have explained that certain thermal effects in the mesoscale ocean have historically been modelled by the thermal quasi-geostrophic (TQG) equations. TQG is characterised by several dimensionless numbers arising from the dimensional parameters of planetary rotation, gravity and buoyancy. These are the familiar Rossby number, Froude number and stratification parameter. The Rossby number is the ratio of a typical horizontal velocity divided by the product of the rotation frequency and a typical horizontal length scale. The Froude number is the ratio of a typical horizontal velocity divided by the velocity of the fastest propagating gravity wave, which in turn is given by the square root of the gravity times the typical vertical length scale. The final dimensionless number is the stratification parameter, which specifies the typical size of the buoyancy stratification.

The regime in which the thermal quasi-geostrophic equations are derived is characterised by the thermal geostrophic balance. This balance arises because the Rossby number, the Froude number and the stratification parameter all have a similar amplitude. Preserving this threefold balance requires simultaneously adapting the Froude number and the stratification parameter to match any change in Rossby number, for example, so that the dimensionless parameters will still have the same size. We consider the non-dissipative case, because of the large scales of mesoscale ocean dynamics. As mentioned earlier, the mesoscale dynamics has high wavenumber instabilities, which in principle can generate submesoscale effects. At smaller scales, viscous dissipation

and thermal diffusivity will come into play as well. However, in what follows, dissipative effects will be neglected. The derivation of the TQG model involves a series of coordinated approximations to the rotating, stratified Euler model, as is illustrated in the diagram below.

The derivation of the right-most three columns of the tree in Fig. 3 is discussed in detail (Holm and Luesink 2021). The derivation of the thermal quasi-geostrophic equations treated here from the thermal rotating shallow water equations on the middle left of the figure is discussed in Holm et al. (2021). For the physical and mathematical details of the derivation indicated by the solid red arrows in Fig. 3, we refer to Holm and Luesink (2021); Holm et al. (2021). The following is a short summary of the derivation indicated by the solid red arrows. The starting point is the Lagrangian for the rotating, stratified Euler equations at the top of the figure. After identifying the small dimensionless parameters in the Euler Lagrangian, the Lagrangians for the successive approximate models can be derived by inserting asymptotic expansions into the Euler Lagrangian. In the ocean, the buoyancy stratification is typically small. Hence, it makes sense to apply the Boussinesq approximation in the Lagrangian for the rotating, stratified Euler equations. This approximation yields the Lagrangian for the Euler–Boussinesq equations. At this point, one has a choice of two routes. The first route begins by making the hydrostatic approximation, which leads to the Lagrangian for the primitive equations. Then, upon vertically integrating the Lagrangian for the primitive equations, one finds the Lagrangian for the thermal rotating shallow water equations. The alternative route first vertically integrates the Lagrangian for the Euler–Boussinesq equations to obtain the Lagrangian for the thermal and rotating version of the Green–Naghdi equations. By subsequently making the hydrostatic approximation in the Lagrangian for the thermal rotating Green–Naghdi equations, the alternative route arrives at the Lagrangian for the thermal rotating shallow water equations.

The Lagrangian for the thermal rotating shallow water (TRSW) equations is the starting point in Holm et al. (2021) for the derivation of the thermal quasi-geostrophic (TQG) equations. The typical values of the dimensionless numbers for mesoscale ocean problems lead to thermal geostrophic balance. This balance implies an algebraic expression for the balanced velocity field in terms of the horizontal gradients of the free surface elevation and the buoyancy. Upon expanding the TRSW Lagrangian around this balance and truncating, one obtains the Lagrangian for the thermal Lagrangian 1 (L1) model. The Lagrangian for the thermal L1 model is not hyperregular, though. Hence, the Legendre transformation for thermal L1 is not invertible. Hence, no Hamiltonian description of this model is available via the Legendre transformation. Nonetheless, an application of the Euler–Poincaré theorem (Holm et al. 1998a) to the thermal L1 Lagrangian yields the corresponding equations of motion. By identifying the leading order terms in the thermal L1 equations and truncating the asymptotic expansion, one finally obtains the TQG equations. However, the truncations of the asymptotic expansions in the thermal L1 equations also prevent the resulting TQG equations from possessing a Hamilton’s principle derivation. This feature is unlike the other models in Fig. 3, which all arise from approximated Lagrangians in their corresponding action integrals for Hamilton’s principle. However, as it turns out, the thermal quasi-geostrophic equations do possess a Hamiltonian formulation in terms of a non-standard Lie–Poisson bracket. The Lie–Poisson bracket for the quasi-geostrophic

equations can be obtained via a linear change of variables from the usual semi-direct product Lie–Poisson bracket for fluids. The resulting Hamiltonian formulation for the thermal quasi-geostrophic equations is important for the application of stochastic advection by Lie transport (SALT), introduced in Holm (2015), which requires either a Lagrangian, or a Hamiltonian interpretation of the equations of motion. The analysis of the stochastic thermal quasi-geostrophic equation has been covered in the companion paper (Crisan et al. 2022).

By applying the method of Stochastic Advection by Lie Transport (SALT) to the Hamiltonian formulation of the TQG equations, one obtains a stochastic version of them which preserves the infinite family of integral conserved quantities. More details on the stochastic version can be found in the conclusion section as well as in Holm et al. (2021). A discussion of the remaining models in Fig. 3 can be found in Holm and Luesink (2021).

The TQG model on the two-dimensional flat torus \mathbb{T}^2 can be formulated in two equivalent ways. The first formulation is the advective formulation, in which the equations of motion are given by

$$\frac{\partial}{\partial t} b + \mathbf{u} \cdot \nabla b = 0, \tag{1.1}$$

$$\frac{\partial}{\partial t} \omega + \mathbf{u} \cdot \nabla (\omega - b) = -\mathbf{u}_{h_1} \cdot \nabla b, \tag{1.2}$$

$$\omega = (\Delta - 1)\psi + f_1, \quad \nabla \cdot \mathbf{u} = 0. \tag{1.3}$$

Here b is the vertically averaged buoyancy, \mathbf{u} is the thermal geostrophically balanced velocity field, ω is the potential vorticity, ψ is the streamfunction, h_1 is the spatial variation around a constant bathymetry profile, and f_1 is the spatial variation around a constant background rotation rate. The velocity \mathbf{u} and the streamfunction ψ are related by the equation

$$\mathbf{u} = \nabla^\perp \psi, \tag{1.4}$$

where $\nabla^\perp = (-\partial_y, \partial_x)$. The vector field \mathbf{u}_{h_1} is defined by

$$\mathbf{u}_{h_1} := \frac{1}{2} \nabla^\perp h_1. \tag{1.5}$$

Alternatively, one can formulate the thermal quasi-geostrophic (TQG) equations in vorticity-streamfunction form. This formulation is given by

$$\frac{\partial}{\partial t} b + J(\psi, b) = 0, \tag{1.6}$$

$$\frac{\partial}{\partial t} \omega + J(\psi, \omega - b) = -\frac{1}{2} J(h_1, b), \tag{1.7}$$

$$\omega = (\Delta - 1)\psi + f_1. \tag{1.8}$$

The operator $J(a, b) = \nabla^\perp a \cdot \nabla b = a_x b_y - b_x a_y$ is the Jacobian of two smooth functions a and b defined on the (x, y) plane. The equivalence of the two formulations

(1.1)–(1.2) and (1.6)–(1.7) follows from the Jacobian operator relation $J(\psi, a) = \mathbf{u} \cdot \nabla a$, in which ψ is the streamfunction associated to the velocity vector field \mathbf{u} . The scalar functions f_1 and h_1 relate to the usual Coriolis parameter and bathymetry profile in the following way

$$\begin{aligned} h(\mathbf{x}) &= 1 + \text{Ro} h_1(\mathbf{x}), \\ f(\mathbf{x}) &= 1 + \text{Ro} f_1(\mathbf{x}), \end{aligned} \quad (1.9)$$

where $\text{Ro} = U(f_0 L)^{-1}$ is the Rossby number, expressed in terms of the typical horizontal velocity U , typical rotation frequency f_0 and typical horizontal length scale L . This means that the bathymetry and Coriolis parameter become constant as the Rossby number tends to zero. Equation (1.9) are necessary to derive the thermal quasi-geostrophic equations from the thermal L1 equations, as shown in Holm et al. (2021). The expansion (1.9) contains the β -plane approximation provided that the boundary conditions are appropriate. Namely, on the β -plane, one requires that $\beta f_0^{-1} = \mathcal{O}(\text{Ro})$ and $f_1(\mathbf{x}) = y$. An additional relation can be helpful when using the thermal quasi-geostrophic equations as a model for mesoscale ocean dynamics. In particular, the streamfunction is related to the free surface elevation ζ and buoyancy b via the definition $\psi := \zeta + \frac{1}{2}b$. This definition of the streamfunction is useful to relate to observational data. The free surface elevation is a quantity that can be measured with satellite altimetry. These measurements can then be used for data assimilation and model calibration. However, the definition of the streamfunction in terms of the velocity field is not necessary to formulate the model as a closed set of equations, since the system (1.6)–(1.8) is already a closed set of equations. In the expansion (1.9), we will henceforth drop the subscript 1 on $h_1(\mathbf{x})$ and $f_1(\mathbf{x})$ for notational convenience.

1.4 Plan for the Rest of the Paper

We now give the plan for the rest of the paper. We collect preliminary tools in Sect. 2. This includes notations and analytical properties of function spaces used throughout this paper. We also collect useful estimates that will be used at various stages and give precise definitions of the concept of solutions used in our analysis. We finally end Sect. 2 with statements of the main results. In particular, we state that both the TQG and α -TQG equations admit unique strong solutions for a finite period of time and these solutions are stable in a larger space with weaker norm. Furthermore, a maximum time for these solutions exists.

Since the proof of local well-posedness is the same for the TQG and α -TQG, we will avoid duplication by devoting Sect. 3 to the construction of solutions for the less regular TQG. The construction relies heavily on the standard energy method. Since we are constructing strong solutions (rather than weak ones), we differentiate the equations in space and then we test the resulting equations with the required differential of the solution to obtain the required bounds. We then end Sect. 3 by showing that the unique solution constructed has a maximum time of existence and hence, is a maximal solution.

In Sect. 4 we show that any family of maximal solutions of the α -TQG models converges strongly with $\alpha \rightarrow 0$ to the unique maximal solution of the TQG on a common existence time, provided they share the same data.

Next, since our solutions are local in nature, we establish in Sect. 5, conditions under which this solution may blow up in the sense of Beale et al. (1984). In particular, we show that in order to control the solution of α -TQG, the essential supremum in space of both the buoyancy gradient and the velocity gradient should be integrable over the anticipated time interval. Once either of these gradients blow up, the solution ceases to exist. Alternatively, in order to control the solution, it suffices to control the H^1 -Sobolev norm of the buoyancy gradient.

Section 6 is devoted to numerical methods and simulation results. We begin by describing the finite element method we use for the spatial derivatives, and aspects of its numerical conservation properties with respect to theoretical results. We then describe the finite difference discretisation method used for the time derivative. Next, we discuss α -TQG linear thermal Rossby wave stability analysis, which can be seen as generalising the results shown in Holm et al. (2021) for the TQG system. Then in the last part of the subsection, we discuss our numerical simulation setup and its results. In particular, we numerically verify the theoretical convergence rate for α -TQG \rightarrow TQG derived in Sect. 4.

2 Preliminaries and Main Results

In this section, we fix the notation, collect some preliminary material on function spaces and present the main analytical results.

Remark 2.1 Although we will be working on the two-dimensional torus, with minimal effort, the same analysis will work on the whole plane \mathbb{R}^2 subject to a far-field condition. The case of a bounded domain with boundary conditions is however outside the scope of the analytical aspect of this work. See Sect. 6 for numerical works in this regard.

2.1 Notations

Our independent variables consists of spatial points $x := \mathbf{x} = (x, y) \in \mathbb{T}^2$ on the 2-torus \mathbb{T}^2 and a time variable $t \in [0, T]$ where $T > 0$. For functions F and G , we write $F \lesssim G$ if there exists a generic constant $c > 0$ such that $F \leq cG$. We also write $F \lesssim_p G$ if the constant $c(p) > 0$ depends on a variable p . The symbol $|\cdot|$ may be used in four different contexts. For a scalar function $f \in \mathbb{R}$, $|f|$ denotes the absolute value of f . For a vector $\mathbf{f} \in \mathbb{R}^2$, $|\mathbf{f}|$ denotes the Euclidean norm of \mathbf{f} . For a square matrix $\mathbb{F} \in \mathbb{R}^{2 \times 2}$, $|\mathbb{F}|$ shall denote the Frobenius norm $\sqrt{\text{trace}(\mathbb{F}^T \mathbb{F})}$. Finally, if $S \subseteq \mathbb{R}^2$ is a (sub)set, then $|S|$ is the two-dimensional Lebesgue measure of S .

For $k \in \mathbb{N} \cup \{0\}$ and $p \in [1, \infty]$, we denote by $W^{k,p}(\mathbb{T}^2)$, the Sobolev space of Lebesgue measurable functions whose weak derivatives up to order k belongs to $L^p(\mathbb{T}^2)$. Its associated norm is

$$\|v\|_{W^{k,p}(\mathbb{T}^2)} = \sum_{|\beta| \leq k} \|\partial^\beta v\|_{L^p(\mathbb{T}^2)}, \tag{2.1}$$

where β is a 2-tuple multi-index of non-negative integers of length $|\beta| \leq k$. The Sobolev space $W^{k,p}(\mathbb{T}^2)$ is a Banach space. Moreover, $W^{k,2}(\mathbb{T}^2)$ is a Hilbert space when endowed with the inner product

$$\langle u, v \rangle_{W^{k,2}(\mathbb{T}^2)} = \sum_{|\beta| \leq k} \langle \partial^\beta u, \partial^\beta v \rangle, \tag{2.2}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard L^2 -inner product. In general, for $s \in \mathbb{R}$, we will define the Sobolev space $H^s(\mathbb{T}^2)$ as consisting of distributions v defined on \mathbb{T}^2 for which the norm

$$\|v\|_{H^s(\mathbb{T}^2)} = \left(\sum_{\xi \in \mathbb{Z}^2} (1 + |\xi|^2)^s |\widehat{v}(\xi)|^2 \right)^{\frac{1}{2}} \equiv \|v\|_{W^{s,2}(\mathbb{T}^2)} \tag{2.3}$$

defined in frequency space is finite. Here, $\widehat{v}(\xi)$ denotes the Fourier coefficients of v . To shorten notation, we will write $\|\cdot\|_{s,2}$ for $\|\cdot\|_{W^{s,2}(\mathbb{T}^2)}$ and/or $\|\cdot\|_{H^s(\mathbb{T}^2)}$. When $k = s = 0$, we get the usual $L^2(\mathbb{T}^2)$ space whose norm we will denote by $\|\cdot\|_2$ for simplicity. We will also use a similar convention for norms $\|\cdot\|_p$ of general $L^p(\mathbb{T}^2)$ spaces for any $p \in [1, \infty]$ as well as for the inner product $\langle \cdot, \cdot \rangle_{k,2} := \langle \cdot, \cdot \rangle_{W^{k,2}(\mathbb{T}^2)}$ when $k \in \mathbb{N}$. Additionally, we will denote by $W_{\text{div}}^{k,p}(\mathbb{T}^2; \mathbb{R}^2)$, the space of weakly divergence-free vector-valued functions in $W^{k,p}(\mathbb{T}^2)$. Finally, we define the following space

$$\mathcal{M} = W^{3,2}(\mathbb{T}^2) \times W^{2,2}(\mathbb{T}^2) \tag{2.4}$$

endowed with the norm

$$\|(b, \omega)\|_{\mathcal{M}} := \|b\|_{3,2} + \|\omega\|_{2,2}. \tag{2.5}$$

2.2 Preliminary Estimates

We begin this section with the following result which follows from a direct computation using the definition (2.3) of the Sobolev norms.

Lemma 2.2 *Let $k \in \mathbb{N} \cup \{0\}$ and assume that the triple $(\psi, \mathbf{u}, \omega)$ satisfies*

$$\mathbf{u} = \nabla^\perp \psi, \quad \omega = (\Delta - 1)\psi. \tag{2.6}$$

If $\omega \in W^{k,2}(\mathbb{T}^2)$, then the following estimate

$$\|\mathbf{u}\|_{k+1,2}^2 \lesssim \|\omega\|_{k,2}^2 \tag{2.7}$$

holds.

Proof Because $\mathbf{u} = \nabla^\perp(\Delta - 1)^{-1}\omega$, it follows that

$$\begin{aligned} \|\mathbf{u}\|_{k+1,2}^2 &= \sum_{\xi \in \mathbb{Z}^2} \frac{(1 + |\xi|^2)^{k+1} |\xi|^2}{(1 + |\xi|^2)^2} |\omega|^2 \\ &\lesssim \sum_{\xi \in \mathbb{Z}^2} \frac{(1 + |\xi|^2)^{k+1} (1 + |\xi|^2)}{(1 + |\xi|^2)^2} |\omega|^2 = \|\omega\|_{k,2}^2. \end{aligned}$$

□

Let us now recall some Moser-type calculus. See (Klainerman and Majda 1981; Kato 1990; Kato and Ponce 1988).

Lemma 2.3 (Commutator estimates) *Let β be a 2-tuple multi-index of non-negative integers such that $|\beta| \leq k$ holds for $k \in \{1, 2\}$. Let $p, p_2, p_3 \in (1, \infty)$ and $p_1, p_4 \in (1, \infty]$ be such that*

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p_3} + \frac{1}{p_4}.$$

For $u \in W^{k,p_3}(\mathbb{T}^2) \cap W^{1,p_1}(\mathbb{T}^2)$ and $v \in W^{k-1,p_2}(\mathbb{T}^2) \cap L^{p_4}(\mathbb{T}^2)$, we have and

$$\|\partial^\beta(uv) - u\partial^\beta v\|_p \lesssim (\|\nabla u\|_{p_1} \|v\|_{k-1,p_2} + \|u\|_{k,p_3} \|v\|_{p_4}). \tag{2.8}$$

2.3 Main Results

Our current goal is to construct a solution for the system of Eqs. (1.1)–(1.5). To do this, we first make the following assumption on our set of data.

Assumption 2.4 Let $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ and $f \in W^{2,2}(\mathbb{T}^2)$ and assume that $(b_0, \omega_0) \in \mathcal{M}$.

Unless otherwise stated, Assumption 2.4 now holds throughout the rest of the paper. With stronger assumptions on the regularity, similar results will also follow. We are now in a position to make precise, exactly what we mean by a solution.

Definition 2.5 (Local strong solution) Let $T > 0$ be a constant. We call the triple (b, ω, T) a *local strong solution* or simply, a *local solution* or a *solution* to the system (1.1)–(1.5) if the following holds.

- The buoyancy b satisfies $b \in C([0, T]; W^{3,2}(\mathbb{T}^2))$ and the equation

$$b(t) = b_0 - \int_0^t \text{div}(b\mathbf{u}) \, d\tau,$$

holds for all $t \in [0, T]$;

- the potential vorticity ω satisfies $\omega \in C([0, T]; W^{2,2}(\mathbb{T}^2))$ and the equation

$$\omega(t) = \omega_0 - \int_0^t \left[\operatorname{div}((\omega - b)\mathbf{u}) + \operatorname{div}(b\mathbf{u}_h) \right] d\tau$$

holds for all $t \in [0, T]$.

- The Biot–Savart law $\mathbf{u} = \nabla^\perp(\Delta - 1)^{-1}\omega$ holds.

Remark 2.6 We remark that the regularity of the solution (b, ω, T) and its data together with the integral equations immediately imply that b and ω are differentiable in time. Indeed, by using the fact that $W^{3,2}(\mathbb{T}^2)$ and $W^{2,2}(\mathbb{T}^2)$ are Banach algebras, we immediately deduce from the integral equations for b and ω above that

$$b \in C^1([0, T]; W^{2,2}(\mathbb{T}^2)), \quad \omega \in C^1([0, T]; W^{1,2}(\mathbb{T}^2)).$$

It also follows from the integral equations for the buoyancy and potential vorticity above that the initial conditions are $b(0, x) = b_0(x)$ and $\omega(0, x) = \omega_0(x)$. The corresponding differential forms (1.1)–(1.2) are clearly immediate from the integral representations.

Remark 2.7 Since we are working on the torus, and the velocity fields are defined by (1.4)–(1.5), we have in particular, $\int_{\mathbb{T}^2} \mathbf{u}_h \, d\mathbf{x} = 0$ and $\int_{\mathbb{T}^2} \mathbf{u} \, d\mathbf{x} = 0$. From the latter, we get that ω and f have zero averages. Consequently, we will assume that all functions under consideration have zero averages.

Definition 2.8 (Maximal solution) We call (b, ω, T_{\max}) a *maximal solution* to the system (1.1)–(1.5) if:

- there exists an increasing sequence of time steps $(T_n)_{n \in \mathbb{N}}$ whose limit is $T_{\max} \in (0, \infty]$;
- for each $n \in \mathbb{N}$, the triple (b, ω, T_n) is a local strong solution to the system (1.1)–(1.5) with initial condition (b_0, ω_0) ;
- if $T_{\max} < \infty$, then

$$\limsup_{T_n \rightarrow T_{\max}} \|(b, \omega)(T_n)\|_{\mathcal{M}}^2 = \infty. \tag{2.9}$$

We shall call $T_{\max} > 0$ the *maximal time*.

Remark 2.9 Condition (2.9) means that the solution breaks down at the limit point T_{\max} .

We are now in a position to state our first main result.

Theorem 2.10 (Existence of local solutions) *There exists a solution (b, ω, T) of (1.1)–(1.5) under Assumption 2.4.*

Once we have constructed a local solution, we can show that this solution is continuously dependent on its initial state in a more general class of function spaces. This choice of class appears to be the strongest space in which the analysis may be performed. More details will follow in the sequel, but first, we give the statement on the continuity property of strong solutions with respect to its data.

Theorem 2.11 (Continuity at low regularity) *Let $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ and $f \in W^{2,2}(\mathbb{T}^2)$. Assume that (b^1, ω^1, T^1) and (b^2, ω^2, T^2) are solutions of (1.1)–(1.5) with initial conditions $(b_0^1, \omega_0^1) \in \mathcal{M}$ and $(b_0^2, \omega_0^2) \in \mathcal{M}$, respectively. Then there exists a constant*

$$c = c(\|b_0^1\|_{3,2}, \|\omega_0^1\|_{2,2}, \|b_0^2\|_{3,2}, \|\omega_0^2\|_{2,2}, \|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2})$$

such that

$$\|(b^1 - b^2)(t)\|_{2,2}^2 + \|(\omega^1 - \omega^2)(t)\|_{1,2}^2 \leq \exp(cT)(\|b_0^1 - b_0^2\|_{2,2}^2 + \|\omega_0^1 - \omega_0^2\|_{1,2}^2) \tag{2.10}$$

holds for all $t \in [0, T]$ where $T = \min\{T^1, T^2\}$.

Remark 2.12 For the avoidance of doubt, we make clear that the final estimate for the differences in Theorem 2.11 above is stated in terms of a larger space $W^{2,2}(\mathbb{T}^2) \times W^{1,2}(\mathbb{T}^2)$ with smaller norms than the space of existence $W^{3,2}(\mathbb{T}^2) \times W^{2,2}(\mathbb{T}^2)$. In order to obtain this bound however, in particular, we require the initial conditions of one of the solutions to be bounded in the stronger space of existence, i.e. boundedness of $\|b_0^1\|_{3,2}, \|\omega_0^1\|_{2,2}$ rather than in the weaker space for which we obtain our final stability estimate. Furthermore, we also require the potential vorticity ω^2 (at all time of existence) of the second solution to be also bounded in the stronger space of existence $W^{2,2}(\mathbb{T}^2)$. Explicitly, the only flexibility we have has to do with the second buoyancy b^2 for which it appears that we are able to relax to live in the weaker space $W^{2,2}(\mathbb{T}^2)$. Unfortunately, however, because of the highly coupled nature of the system of equations under study, needing $\omega^2 \in W^{2,2}(\mathbb{T}^2)$ for all times automatically requires that $b^2 \in W^{3,2}(\mathbb{T}^2)$ for all times. Subsequently, this means that we require boundedness of the initial condition of the second solution in the stronger space of existence, i.e. boundedness of $\|b_0^2\|_{3,2}, \|\omega_0^2\|_{2,2}$ in addition to that of the first solution. The requirement of needing both pairs of initial conditions to live in a stronger space is in contrast to simpler looking models like the Euler equation where it suffices to require just having one initial condition to have stronger regularity. Unfortunately, we can not do better by our method of proof (which is to derive estimates for equations solved by the differences, i.e. the energy method) but we do not claim that other methods for deriving analogous estimates may not yield better result either.

As a consequence of Theorem 2.11, the following statement about uniqueness of the strong solution is immediate.

Corollary 2.13 (Uniqueness) *Let (b^1, ω^1, T^1) and (b^2, ω^2, T^2) be two solutions of (1.1)–(1.5) under Assumption 2.4. Then the difference $(b^1 - b^2, \omega^1 - \omega^2)$ satisfies the*

equation

$$\|(b^1 - b^2)(t)\|_{2,2}^2 + \|(\omega^1 - \omega^2)(t)\|_{1,2}^2 = 0 \tag{2.11}$$

for all $t \in [0, T]$ where $T = \min\{T^1, T^2\}$.

Finally, we can show that a maximal solution of (1.1)–(1.5), in the sense of Definition 2.8, also exists.

Theorem 2.14 (Existence of maximal solution) *There exist a unique maximal solution (b, ω, T_{\max}) of (1.1)–(1.5) under Assumption 2.4.*

The α -TQG model: In the following, we consider a ‘regularised’ version of (1.1)–(1.5). The regularisation is known as α -regularisation and is of a dispersive type, i.e. the regularisation behaves like a low-pass filter, as explained in Holm et al. (1998b). The main benefit of this type of regularisation is that energy is also preserved by the regularised model, contrary to diffusive regularisations that dissipate energy. Since our notion of a solution to (1.1)–(1.5) involves the pair (b, ω) (and not explicitly in terms of \mathbf{u}), henceforth, we will use the triple $(b^\alpha, \omega^\alpha, T)$ to denote the corresponding solution to the following α -TQG model for the avoidance of confusion. To be precise, for fixed $\alpha > 0$, we will be exploring the pair $(b^\alpha, \omega^\alpha)$ that solves

$$\frac{\partial}{\partial t} b^\alpha + \mathbf{u}^\alpha \cdot \nabla b^\alpha = 0, \tag{2.12}$$

$$\frac{\partial}{\partial t} \omega^\alpha + \mathbf{u}^\alpha \cdot \nabla(\omega^\alpha - b^\alpha) = -\mathbf{u}_h \cdot \nabla b^\alpha \tag{2.13}$$

but where now,

$$\mathbf{u}^\alpha = \nabla^\perp \psi^\alpha, \quad \mathbf{u}_h = \frac{1}{2} \nabla^\perp h, \quad \omega^\alpha = (\Delta - 1)(1 - \alpha \Delta) \psi^\alpha + f. \tag{2.14}$$

Note that at least formally, this implies that

$$\mathbf{u}^\alpha = \nabla^\perp (1 - \alpha \Delta)^{-1} (\Delta - 1)^{-1} (\omega^\alpha - f). \tag{2.15}$$

We also note that even though the velocity fields \mathbf{u} in (1.5) and \mathbf{u}^α in (2.15) are different, the coupled equations (1.1)–(1.2) and (2.12)–(2.13) are exactly of the same form. In the following, we claim that the exact same result, Theorem 2.14 holds for the α -TQG model (2.12)–(2.14) introduced above.

Theorem 2.15 *Fix $\alpha > 0$. There exists a unique maximal solution $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)$ of (2.12)–(2.14) under Assumption 2.4.*

Remark 2.16 The fact that Theorem 2.15 holds true is hardly surprising considering that we are basically treating the exact same model as (1.1)–(1.2) albeit different velocity fields. Heuristically, one observes that for very small α , $\psi^\alpha \sim \tilde{\psi} \sim \psi$ and thus, both models are the same. Nevertheless, rigorously, the key lies in the estimate

(2.7) which still holds true when one replaces \mathbf{u} with \mathbf{u}^α and where now, the triple $(\mathbf{u}^\alpha, \omega^\alpha, f)$ satisfies

$$\mathbf{u}^\alpha = \nabla^\perp \psi^\alpha, \quad (\omega^\alpha - f) = (\Delta - 1)(1 - \alpha\Delta)\psi^\alpha. \tag{2.16}$$

for a given f having sufficient regularity.

Indeed, Theorem 2.15 follows from the following lemma.

Lemma 2.17 *Under Assumption 2.4, there exists a unique solution $(b^\alpha, \omega^\alpha, T^\alpha)$ of (2.12)–(2.14) for any $\alpha > 0$. Moreover, $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)$ is a unique maximal solution of (2.12)–(2.14).*

Proof The existence of a unique solution $(b^\alpha, \omega^\alpha, T^\alpha)$ of (2.12)–(2.14) for any $\alpha > 0$ follows from Proposition 3.1, Lemma 3.5 and Corollary 2.13 with $T^\alpha = T$. We will prove all these results later for the TQG and one will observe that nothing changes for the α -TQG. Note that since (1.1)–(1.5) and (2.12)–(2.14) share the same data, Assumption 2.4, the local existence time T is accordingly independent of $\alpha > 0$.

The construction of the maximal solution $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)$ of (2.12)–(2.14) also follows the same gluing argument used in showing the existence of the maximal solution (b, ω, T_{\max}) for (1.1)–(1.5) in Sect. 3.4. □

Remark 2.18 Even though we have no proof at the moment, we expect that the solution of the α -TQG model (2.12)–(2.14) is in fact smooth (perhaps even analytical), this is not expected to hold globally in time. In other words, the α -TQG model is expected to be a regularisation of the TQG model, as it limits the rate at which the energy can go to high wave numbers. This is clear from the growth rates for linear stability analysis, see Sect. 6.2 for details.

3 Construction of Strong Solutions

In this section, we give a construction of a solution, in the sense of Definition 2.5, to the thermal quasi-geostrophic system of Eqs. (1.1)–(1.5). We will achieve this goal by rewriting our set of Eqs. (1.1)–(1.5) in an abstract form and show that the resulting operator acting on the pair (b, ω) satisfies the assumptions of Theorem 8.1 in the appendix. The details are as follows. We consider

$$\frac{\partial}{\partial t} \begin{pmatrix} b \\ \omega \end{pmatrix} + \mathcal{A} \begin{pmatrix} b \\ \omega \end{pmatrix} = 0, \quad t \geq 0, \quad b|_{t=0} = b_0, \quad \omega|_{t=0} = \omega_0 \tag{3.1}$$

where the operator

$$\mathcal{A} := \begin{bmatrix} \mathbf{u} \cdot \nabla & 0 \\ (\mathbf{u}_h - \mathbf{u}) \cdot \nabla & \mathbf{u} \cdot \nabla \end{bmatrix} \tag{3.2}$$

is defined such that

$$\mathbf{u} = \nabla^\perp \psi, \quad \mathbf{u}_h = \frac{1}{2} \nabla^\perp h, \quad \omega = (\Delta - 1)\psi + f. \tag{3.3}$$

3.1 Estimates for Convective Terms

In the following, we let $k \in \mathbb{N} \cup \{0\}$ and recall that $W^{k,p}(\mathbb{T}^2)$ is a Banach algebra when $kp > 2$. By using Sobolev embeddings, the Biot–Savart law and (2.7), we obtain the following estimates.

(1b) Let $k \in \{0, 1, 2\}$. If $b \in W^{k+1,2}(\mathbb{T}^2)$, $\omega \in W^{2,2}(\mathbb{T}^2)$ and $\mathbf{u} \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ solves (1.3), (1.4), (1.5) for a given $f \in W^{2,2}(\mathbb{T}^2)$ and a given $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$, then

$$\|\mathbf{u} \cdot \nabla b\|_{k,2} \lesssim \|\mathbf{u}\|_{2,2} \|b\|_{k+1,2} \lesssim (1 + \|\omega\|_{1,2}) \|b\|_{k+1,2}$$

and

$$\begin{aligned} \left| \langle \mathbf{u} \cdot \nabla b, b \rangle_{k,2} \right| &\lesssim \|\mathbf{u}\|_{3,2} \|b\|_{k,2}^2 \lesssim (1 + \|\omega\|_{2,2}) \|b\|_{k,2}^2, \\ \left| \langle \mathbf{u} \cdot \nabla b, \omega \rangle_{k,2} \right| &\lesssim \|\mathbf{u}\|_{2,2} \|b\|_{k+1,2} \|\omega\|_{k,2} \lesssim (1 + \|\omega\|_{1,2}) \|b\|_{k+1,2} \|\omega\|_{k,2}, \\ \left| \langle \mathbf{u}_h \cdot \nabla b, \omega \rangle_{k,2} \right| &\lesssim \|\mathbf{u}_h\|_{2,2} \|b\|_{k+1,2} \|\omega\|_{k,2} \lesssim \|b\|_{k+1,2} \|\omega\|_{k,2}. \end{aligned}$$

(1b) Let $k \geq 3$. If $b \in W^{k+1,2}(\mathbb{T}^2)$, $\omega \in W^{k,2}(\mathbb{T}^2)$ and $\mathbf{u} \in W_{\text{div}}^{k+1,2}(\mathbb{T}^2; \mathbb{R}^2)$ solves (1.3), (1.4), (1.5) for a given $f \in W^{k,2}(\mathbb{T}^2)$ and a given $\mathbf{u}_h \in W_{\text{div}}^{k+1,2}(\mathbb{T}^2; \mathbb{R}^2)$, then

$$\|\mathbf{u} \cdot \nabla b\|_{k,2} \lesssim \|\mathbf{u}\|_{2,2} \|b\|_{k+1,2} + \|\mathbf{u}\|_{k,2} \|b\|_{3,2} \lesssim (1 + \|\omega\|_{k-1,2}) \|b\|_{k+1,2}$$

and

$$\begin{aligned} \left| \langle \mathbf{u} \cdot \nabla b, b \rangle_{k,2} \right| &\lesssim \|\mathbf{u}\|_{3,2} \|b\|_{k,2}^2 + \|\mathbf{u}\|_{k,2} \|b\|_{k,2} \|b\|_{3,2} \lesssim (1 + \|\omega\|_{k-1,2}) \|b\|_{k,2}^2, \\ \left| \langle \mathbf{u} \cdot \nabla b, \omega \rangle_{k,2} \right| &\lesssim (\|\mathbf{u}\|_{2,2} \|b\|_{k+1,2} + \|\mathbf{u}\|_{k,2} \|b\|_{3,2}) \|\omega\|_{k,2} \lesssim (1 + \|\omega\|_{k-1,2}) \|b\|_{k+1,2} \|\omega\|_{k,2}, \\ \left| \langle \mathbf{u}_h \cdot \nabla b, \omega \rangle_{k,2} \right| &\lesssim (\|\mathbf{u}_h\|_{2,2} \|b\|_{k+1,2} + \|\mathbf{u}_h\|_{k,2} \|b\|_{3,2}) \|\omega\|_{k,2} \lesssim \|b\|_{k+1,2} \|\omega\|_{k,2}. \end{aligned}$$

(2a) Let $k \in \{0, 1, 2\}$. If $\omega \in W^{k+1,2}(\mathbb{T}^2)$ and that $\mathbf{u} \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ solves (1.3), (1.4), (1.5) for a given $f \in W^{2,2}(\mathbb{T}^2)$, then

$$\|\mathbf{u} \cdot \nabla \omega\|_{k,2} \lesssim \|\mathbf{u}\|_{2,2} \|\omega\|_{k+1,2} \lesssim (1 + \|\omega\|_{1,2}) \|\omega\|_{k+1,2}$$

and

$$\left| \langle \mathbf{u} \cdot \nabla \omega, \omega \rangle_{k,2} \right| \lesssim \|\mathbf{u}\|_{3,2} \|\omega\|_{k,2}^2 \lesssim (1 + \|\omega\|_{2,2}) \|\omega\|_{k,2}^2.$$

(2b) Let $k \geq 3$. If $\omega \in W^{k+1,2}(\mathbb{T}^2)$ and $\mathbf{u} \in W_{\text{div}}^{k+1,2}(\mathbb{T}^2; \mathbb{R}^2)$ solves (1.3), (1.4), (1.5) for a given $f \in W^{k,2}(\mathbb{T}^2)$, then

$$\|\mathbf{u} \cdot \nabla \omega\|_{k,2} \lesssim \|\mathbf{u}\|_{2,2} \|\omega\|_{k+1,2} + \|\mathbf{u}\|_{k,2} \|\omega\|_{3,2} \lesssim (1 + \|\omega\|_{k-1,2}) \|\omega\|_{k+1,2}$$

and

$$\left| \langle \mathbf{u} \cdot \nabla \omega, \omega \rangle_{k,2} \right| \lesssim \|\mathbf{u}\|_{3,2} \|\omega\|_{k,2}^2 + \|\mathbf{u}\|_{k,2} \|\omega\|_{k,2} \|\omega\|_{3,2} \lesssim (1 + \|\omega\|_{k-1,2}) \|\omega\|_{k,2}^2.$$

With these estimates in hand, we can now proceed to prove the existence of a local strong solution of (3.1).

3.2 Proof of Local Existence

Our proof of a local strong solution to (1.1)–(1.5) will follow from the following proposition. Its proof will follow the argument presented in Kato and Lai (1984) (see Appendix, Sect. 8) for the construction of a local solution to the Euler equation.

Proposition 3.1 *Take Assumption 2.4. There exists a solution (b, ω, T) of (3.1) such that:*

(1) *the time $T > 0$ satisfy the bound*

$$T < \frac{1}{2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)}, \quad c = c(\|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2}); \tag{3.4}$$

(2) *the pair (b, ω) is of class*

$$b \in C_w([0, T]; W^{3,2}(\mathbb{T}^2)), \quad \omega \in C_w([0, T]; W^{2,2}(\mathbb{T}^2)); \tag{3.5}$$

(3) *the pair (b, ω) satisfies the bound*

$$\|(b, \omega)(t)\|_{\mathcal{M}}^2 \leq \frac{\|(b_0, \omega_0)\|_{\mathcal{M}}^2 + 2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2) t}{1 - 2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2) t}, \quad t \in (0, T). \tag{3.6}$$

Proof Let us consider the following triplet $\{V, H, X\}$ given as follows. Take $X := W^{1,2}(\mathbb{T}^2) \times L^2(\mathbb{T}^2)$ and let $H := W^{(3,2),2}(\mathbb{T}^2) \times W^{(2,1),2}(\mathbb{T}^2)$ be a Hilbert space endowed with the inner product

$$\langle (b, \omega), (b', \omega') \rangle_H = \langle b, b' \rangle_{3,2} + \langle b, b' \rangle_{2,2} + \langle \omega, \omega' \rangle_{2,2} + \langle \omega, \omega' \rangle_{1,2}. \tag{3.7}$$

We note that for $(b, \omega) \in \mathcal{M}$, we have that

$$\|(b, \omega)\|_{\mathcal{M}} \leq \|(b, \omega)\|_H \leq 2\|(b, \omega)\|_{\mathcal{M}} \tag{3.8}$$

and thus, H is equivalent to \mathcal{M} . Also, by virtue of the estimates shown in Sect. 3.1, we can conclude that the operator \mathcal{A} is weakly continuous from H into X .

Next, we let V be the domain of an unbounded selfadjoint operator $S \geq 0$ in X with $\text{domain}(S) \subset H$ and $\langle Su, v \rangle = \langle u, v \rangle_H$ for $u \in \text{domain}(S)$. More precisely, $S = \sum_{|\beta| \leq 3} (-\partial)^\beta \partial^\beta$ subject to periodic boundary conditions. With this definition of V in hand, we can again conclude from the estimates in Sect. 3.1 that \mathcal{A} maps V into H . Our goal now is to show that for any $(b, \omega) \in V$, there exists an increasing function $\rho(r) \geq 0$ of $r \geq 0$ such that

$$\left| \langle \mathcal{A}(b, \omega), (b, \omega) \rangle_H \right| \leq \rho(\|(b, \omega)\|_H^2). \tag{3.9}$$

To achieve this goal, we further refine the following estimates from Sect. 3.1. In particular, we recall that if $b \in W^{k+1,2}(\mathbb{T}^2)$ and $\omega \in W^{2,2}(\mathbb{T}^2)$ and that $\mathbf{u} \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ solves (1.3), (1.4), (1.5) for a given $f \in W^{2,2}(\mathbb{T}^2)$ and a given $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$, then for $k \in \{0, 1, 2\}$,

$$\begin{aligned} \left| \langle \mathbf{u} \cdot \nabla b, b \rangle_{k,2} \right| &\lesssim (1 + \|\omega\|_{2,2}) \|b\|_{k+1,2}^2 \lesssim (1 + \|(b, \omega)\|_H^2) \|b\|_{3,2}^2, \\ \left| \langle \mathbf{u} \cdot \nabla \omega, \omega \rangle_{k,2} \right| &\lesssim (1 + \|\omega\|_{2,2}) \|\omega\|_{2,2}^2 \lesssim (1 + \|(b, \omega)\|_H^2) \|\omega\|_{2,2}^2, \\ \left| \langle \mathbf{u} \cdot \nabla b, \omega \rangle_{k,2} \right| &\lesssim (1 + \|\omega\|_{2,2}) \|b\|_{k+1,2} \|\omega\|_{2,2} \lesssim (1 + \|(b, \omega)\|_H^2) (\|b\|_{3,2}^2 + \|\omega\|_{2,2}^2), \\ \left| \langle \mathbf{u}_h \cdot \nabla b, \omega \rangle_{k,2} \right| &\lesssim \|b\|_{k+1,2} \|\omega\|_{2,2} \lesssim (\|b\|_{3,2}^2 + \|\omega\|_{2,2}^2) \lesssim (1 + \|(b, \omega)\|_H^2). \end{aligned} \tag{3.10}$$

In the above estimates, we have used inequalities such as $x \lesssim 1 + x^2$ for $x \geq 0$ and Young’s inequalities. Furthermore, if $b \in W^{3,2}(\mathbb{T}^2)$, $\omega \in W^{2,2}(\mathbb{T}^2)$ and $\mathbf{u} \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$, we also have that

$$\left| \langle \mathbf{u} \cdot \nabla b, b \rangle_{3,2} \right| \lesssim (1 + \|\omega\|_{2,2}) \|b\|_{3,2}^2 \lesssim (1 + \|(b, \omega)\|_H^2) \|b\|_{3,2}^2. \tag{3.11}$$

We can therefore conclude that for $(b, \omega) \in H$ (and in particular, for any $(b, \omega) \in V$),

$$\left| \langle \mathcal{A}(b, \omega), (b, \omega) \rangle_H \right| \lesssim (1 + \|(b, \omega)\|_H^2)^2 \tag{3.12}$$

with a constant depending only on $\|\mathbf{u}_h\|_{3,2}$ and $\|f\|_{2,2}$. Since $\rho(r) = c(1 + r)^2 \geq 0$ is an increasing function of $r \geq 0$, we can conclude from the result in the Appendix, Sect. 8, that for a given $(b_0, \omega_0) \in H \equiv \mathcal{M}$, there exists a solution (b, ω, T) of (3.1) of class

$$b \in C_w([0, T]; W^{3,2}(\mathbb{T}^2)), \quad \omega \in C_w([0, T]; W^{2,2}(\mathbb{T}^2)) \tag{3.13}$$

satisfying the bound (recall (3.8))

$$\|(b, \omega)(t)\|_{\mathcal{M}}^2 \leq \|(b, \omega)(t)\|_H^2 \leq r(t), \quad t \in (0, T) \tag{3.14}$$

where r is an increasing function on $(0, T)$. The time $T > 0$ in (3.13)–(3.14) depends only on $\|\mathbf{u}_h\|_{3,2}$ and $\|f\|_{2,2}$ by way of the function $\rho(\cdot)$ as well as on $\|(b_0, \omega_0)\|_{\mathcal{M}}$. To be precise, for $\rho(r) = c(1+r)^2 \geq 0$ where $c > 0$ depends only on $\|\mathbf{u}_h\|_{3,2}$ and $\|f\|_{2,2}$, we obtain $T > 0$ by solving the equation

$$\frac{d}{dt}r = 2\rho(r), \quad r(0) = \|(b_0, \omega_0)\|_{\mathcal{M}}^2 \tag{3.15}$$

which yields

$$r(t) = \frac{\|(b_0, \omega_0)\|_{\mathcal{M}}^2 + 2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2) t}{1 - 2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2) t}. \tag{3.16}$$

We therefore require

$$T < \frac{1}{2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)} \tag{3.17}$$

for a solution of (3.15) to exist. This finishes the proof. □

Remark 3.2 Note that $r(t)$ converges to $\|(b_0, \omega_0)\|_{\mathcal{M}}^2$ as $t \rightarrow 0^+$. Therefore, given the bound (3.14) and (3.16), we can conclude that

$$\limsup_{t \rightarrow 0^+} \|(b, \omega)(t)\|_{\mathcal{M}}^2 \leq \|(b_0, \omega_0)\|_{\mathcal{M}}^2.$$

On the other hand, because the pair (b, ω) is of class (3.13), we can also conclude that

$$\liminf_{t \rightarrow 0^+} \|(b, \omega)(t)\|_{\mathcal{M}}^2 \geq \|(b_0, \omega_0)\|_{\mathcal{M}}^2$$

and thus, we obtain strong right continuity at time $t = 0$. Since our system (1.1)–(1.5) is time reversible (see Remark 3.3 below), we also obtain strong left continuity at $t = 0$ so that we are able to conclude that (3.13) is actually strongly continuous at $t = 0$.

Remark 3.3 By time-reversible, we mean that if $b(t, x)$ and $\omega(t, x)$ (with $\omega(t, x)$ related to $\mathbf{u}(t, x)$ by (1.5)) solves (1.1)–(1.5) with data $b_0(x), \omega_0(x), \mathbf{u}_h(x)$ and $f(x)$, then $-b(-t, x)$ and $-\omega(-t, x)$ (with $-\omega(-t, x)$ related to $-\mathbf{u}(-t, x)$ by (1.5)) solves (1.1)–(1.5) with data $-b_0(x), -\omega_0(x), -\mathbf{u}_h(x)$ and $-f(x)$, respectively.

Remark 3.4 We remark that the function ρ and, as such, the function r solving (3.15) are not unique. In fact, one can construct countably many (if not infinitely many) of these functions by varying the estimates for the convective terms on the left-hand sides

of (3.10)–(3.11). For example, a different choice of Hölder conjugates for estimating the aforementioned terms will lead to a different ρ and r . The importance of this remark lies in the fact that ρ and r determines the longevity of the solution (b, ω, T) to (1.1)–(1.5). One can therefore tune $T > 0$ by modifying ρ and r accordingly. For example, we may also obtain the following

$$\rho_1(r) = c(r^2 + r), \quad \rho_2(r) = c(r^{\frac{3}{2}} + r) \tag{3.18}$$

for which the corresponding solutions to (3.15) are

$$r_1(t) = \frac{r(0)e^{2ct}}{r(0) + 1 - r(0)e^{2ct}}, \tag{3.19}$$

$$r_2(t) = \frac{s^2 + s \pm 2s\sqrt{s}}{s^2 - 2s + 1}, \quad \text{where } s = s(t) = \frac{r(0)}{(\sqrt{r(0)} + 1)^2} e^{2ct}, \tag{3.20}$$

respectively. Again, the constants depends only on $\|\mathbf{u}_h\|_{3,2}$ and $\|f\|_{2,2}$. The second solution r_2 is unsuitable since for one, it is twofold. The first solution r_1 , however, means that we require a time

$$T_1 < \frac{1}{2c} \ln \left(1 + \frac{1}{\|(b_0, \omega_0)\|_{\mathcal{M}}^2} \right)$$

for a solution of (3.15) to exist. Therefore, if we optimise the constant $c > 0$ so that they are the same throughout and we compare r_1 with r given in (3.16), we are able to conclude that r yields a longer time of existence of a solution (b, ω, T) to (1.1)–(1.5) as compared to r_1 since

$$\frac{1}{2c} \ln \left(1 + \frac{1}{\|(b_0, \omega_0)\|_{\mathcal{M}}^2} \right) \leq \frac{1}{2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)}.$$

In the next section, we show an estimate for the difference of two solutions constructed above after which we are able to strengthen the weak continuity (3.13) to a strong one for all times of existence.

3.3 Difference Estimate

In the following, we let (b^1, ω^1, T) and (b^2, ω^2, T) be two solutions of (1.1)–(1.5) sharing the same data $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$, $f \in W^{2,2}(\mathbb{T}^2)$ and (b_0, ω_0) . So in particular,

$$\sup_{t \in [0, T]} \|b^i(t, \cdot)\|_{3,2} < \infty, \quad \sup_{t \in [0, T]} \|\omega^i(t, \cdot)\|_{2,2} < \infty \tag{3.21}$$

hold for each $i = 1, 2$. We now set $b^{12} := b^1 - b^2, \omega^{12} := \omega^1 - \omega^2$ and $\mathbf{u}^{12} := \mathbf{u}^1 - \mathbf{u}^2$ so that $(b^{12}, \mathbf{u}^{12}, \omega^{12})$ satisfies

$$\frac{\partial}{\partial t} b^{12} + \mathbf{u}^{12} \cdot \nabla b^1 + \mathbf{u}^2 \cdot \nabla b^{12} = 0, \tag{3.22}$$

$$\frac{\partial}{\partial t} \omega^{12} + \mathbf{u}^{12} \cdot \nabla(\omega^1 - b^1) + \mathbf{u}^2 \cdot \nabla(\omega^{12} - b^{12}) = -\mathbf{u}_h \cdot \nabla b^{12}, \tag{3.23}$$

where

$$\mathbf{u}^{12} = \nabla^\perp \psi^{12}, \quad \mathbf{u}_h = \frac{1}{2} \nabla^\perp h, \quad \omega^{12} = (\Delta - 1)\psi^{12}. \tag{3.24}$$

We now show the following stability and uniqueness results in the space $W^{2,2}(\mathbb{T}^2) \times W^{1,2}(\mathbb{T}^2)$ which is larger than the space $W^{3,2}(\mathbb{T}^2) \times W^{2,2}(\mathbb{T}^2)$ of existence of the individual solutions. Indeed, it suffices to show uniqueness in the even larger space $W^{1,2}(\mathbb{T}^2) \times L^2(\mathbb{T}^2)$ but we avoid doing so since we wish to concurrently show the proof of Theorem 2.11 as well.

Proof of Theorem 2.11 and Corollary 2.13 If we apply ∂^β to (3.22) with $|\beta| \leq 2$, we obtain

$$\frac{\partial}{\partial t} \partial^\beta b^{12} + \mathbf{u}^2 \cdot \nabla \partial^\beta b^{12} = S_1 + S_2 - \mathbf{u}^{12} \cdot \nabla \partial^\beta b^1, \tag{3.25}$$

where

$$S_1 := \mathbf{u}^{12} \cdot \partial^\beta \nabla b^1 - \partial^\beta (\mathbf{u}^{12} \cdot \nabla b^1),$$

$$S_2 := \mathbf{u}^2 \cdot \partial^\beta \nabla b^{12} - \partial^\beta (\mathbf{u}^2 \cdot \nabla b^{12})$$

are such that

$$\|S_1\|_2 \lesssim \|\nabla \mathbf{u}^{12}\|_4 \|b^1\|_{2,4} + \|\nabla b^1\|_\infty \|\mathbf{u}^{12}\|_{2,2} \lesssim \|\mathbf{u}^{12}\|_{2,2} \|b^1\|_{3,2}, \tag{3.26}$$

$$\|S_2\|_2 \lesssim \|\nabla \mathbf{u}^2\|_\infty \|b^{12}\|_{2,2} + \|\nabla b^{12}\|_4 \|\mathbf{u}^2\|_{2,4} \lesssim \|\mathbf{u}^2\|_{3,2} \|b^{12}\|_{2,2}. \tag{3.27}$$

Here, we have used the commutator estimate (2.3) and the continuous embeddings $W^{2,2}(\mathbb{T}^2) \hookrightarrow W^{1,4}(\mathbb{T}^2), W^{3,2}(\mathbb{T}^2) \hookrightarrow W^{2,4}(\mathbb{T}^2)$ and $W^{3,2}(\mathbb{T}^2) \hookrightarrow W^{1,\infty}(\mathbb{T}^2)$. Additionally, the following estimate for the L^2 inner product of the last term (3.25) with $\partial^\beta b^{12}$ holds

$$\left| \langle \mathbf{u}^{12} \cdot \nabla \partial^\beta b^1, \partial^\beta b^{12} \rangle \right| \lesssim \|b^1\|_{3,2} \|\mathbf{u}^{12}\|_\infty \|b^{12}\|_{2,2}. \tag{3.28}$$

If we test (3.25) with $\partial^\beta b^{12}$ and sum over $|\beta| \leq 2$, we obtain from the above estimates together with (2.7) and (3.21),

$$\frac{d}{dt} \|b^{12}\|_{2,2}^2 \lesssim \|\omega^{12}\|_{1,2}^2 + \|b^{12}\|_{2,2}^2. \tag{3.29}$$

Next, we apply ∂^β to (3.23) with $|\beta| \leq 1$ to get

$$\begin{aligned} \frac{\partial}{\partial t} \partial^\beta \omega^{12} + \mathbf{u}^2 \cdot \nabla \partial^\beta \omega^{12} &= -\mathbf{u}^{12} \cdot \nabla \partial^\beta (\omega^1 - b^1) + (\mathbf{u}^2 - \mathbf{u}_h) \cdot \nabla \partial^\beta b^{12} \\ &+ S_3 + \dots + S_7 \end{aligned} \tag{3.30}$$

where

$$\begin{aligned} S_3 &:= \mathbf{u}^2 \cdot \partial^\beta \nabla \omega^{12} - \partial^\beta (\mathbf{u}^2 \cdot \nabla \omega^{12}), \\ S_4 &:= \mathbf{u}^{12} \cdot \partial^\beta \nabla \omega^1 - \partial^\beta (\mathbf{u}^{12} \cdot \nabla \omega^1), \\ S_5 &:= -\mathbf{u}^{12} \cdot \partial^\beta \nabla b^1 + \partial^\beta (\mathbf{u}^{12} \cdot \nabla b^1), \\ S_6 &:= -\mathbf{u}^2 \cdot \partial^\beta \nabla b^{12} + \partial^\beta (\mathbf{u}^2 \cdot \nabla b^{12}), \\ S_7 &:= \mathbf{u}_h \cdot \partial^\beta \nabla b^{12} - \partial^\beta (\mathbf{u}_h \cdot \nabla b^{12}) \end{aligned}$$

are such that

$$\|S_3\|_2 \lesssim \|\mathbf{u}^2\|_{1,\infty} \|\omega^{12}\|_{1,2} \lesssim \|\omega^{12}\|_{1,2}, \tag{3.31}$$

$$\|S_4\|_2 \lesssim \|\mathbf{u}^{12}\|_{1,4} \|\omega^1\|_{1,4} \lesssim \|\mathbf{u}^{12}\|_{2,2} \lesssim \|\omega^{12}\|_{1,2}, \tag{3.32}$$

$$\|S_5\|_2 \lesssim \|\mathbf{u}^{12}\|_{1,4} \|b^1\|_{1,4} \lesssim \|\mathbf{u}^{12}\|_{2,2} \lesssim \|\omega^{12}\|_{1,2}, \tag{3.33}$$

$$\|S_6\|_2 \lesssim \|\mathbf{u}^2\|_{1,\infty} \|b^{12}\|_{1,2} \lesssim \|b^{12}\|_{1,2} \lesssim \|b^{12}\|_{2,2}, \tag{3.34}$$

$$\|S_7\|_2 \lesssim \|\mathbf{u}_h\|_{1,\infty} \|b^{12}\|_{1,2} \lesssim \|b^{12}\|_{1,2} \lesssim \|b^{12}\|_{2,2}. \tag{3.35}$$

Also, since $\text{div}(\mathbf{u}^2) = 0$

$$\begin{aligned} \left| \langle \mathbf{u}^2 \cdot \nabla \partial^\beta \omega^{12}, \partial^\beta \omega^{12} \rangle \right| &= 0, \\ \left| \langle \mathbf{u}^{12} \cdot \nabla \partial^\beta (\omega^1 - b^1), \partial^\beta \omega^{12} \rangle \right| &\lesssim \|\omega^1 - b^1\|_{2,2} \|\mathbf{u}^{12}\|_\infty \|\omega^{12}\|_{1,2}, \\ \left| \langle (\mathbf{u}^2 - \mathbf{u}_h) \cdot \nabla \partial^\beta b^{12}, \partial^\beta \omega^{12} \rangle \right| &\lesssim \|\mathbf{u}^2 - \mathbf{u}_h\|_\infty \|b^{12}\|_{2,2} \|\omega^{12}\|_{1,2}. \end{aligned} \tag{3.36}$$

Therefore,

$$\frac{d}{dt} \|\omega^{12}\|_{1,2}^2 \lesssim \|\omega^{12}\|_{1,2}^2 + \|b^{12}\|_{2,2}^2. \tag{3.37}$$

From (3.29) and (3.37), we can conclude from Grönwall’s lemma that

$$\|b^{12}(t, \cdot)\|_{2,2}^2 + \|\omega^{12}(t, \cdot)\|_{1,2}^2 \leq \exp(ct) (\|b_0^{12}\|_{2,2}^2 + \|\omega_0^{12}\|_{1,2}^2) \tag{3.38}$$

for all $t \geq 0$ where the constant $c > 0$ depends on $(\|b^1\|_{3,2}, \|\omega^1\|_{2,2}, \|\omega^2\|_{2,2}, \|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2})$. This finishes the proof of Theorem 2.11. To obtain uniqueness, i.e. Corollary 2.13, we use the fact that $b_0^{12} = 0$ and $\omega_0^{12} = 0$. \square

Having shown uniqueness, we can conclude that the weakly continuous functions (3.5) are indeed strongly continuous by using time-reversibility of (1.1)–(1.2).

Lemma 3.5 *Take the assumptions of Proposition 3.1 to be true. Then*

$$b \in C([0, T]; W^{3,2}(\mathbb{T}^2)), \quad \omega \in C([0, T]; W^{2,2}(\mathbb{T}^2)).$$

Proof Firstly, we recall that the solution we constructed in Proposition 3.1 is actually strongly continuous at time $t = 0$, recall Remark 3.2. Now consider a solution (b, ω, T_0) of (1.1)–(1.2) where $T_0 \in [0, T]$ is fixed but arbitrary. Then by (3.6), this solution will satisfy the bound

$$\|(b, \omega)(T_0)\|_{\mathcal{M}}^2 \leq \frac{\|(b_0, \omega_0)\|_{\mathcal{M}}^2 + 2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)T_0}{1 - 2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)T_0} \tag{3.39}$$

$c = c(\|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2})$. We can now construct a new solution $(b^1, \omega^1, T_0 + T_1)$ by taking $(b, \omega)(T_0)$ as an initial condition. A repetition of the argument leading to (3.39) means that this new pair (b^1, ω^1) solving (1.1)–(1.2) will satisfy the inequality

$$\|(b^1, \omega^1)(t)\|_{\mathcal{M}}^2 \leq \frac{\|(b, \omega)(T_0)\|_{\mathcal{M}}^2 + 2c(1 + \|(b, \omega)(T_0)\|_{\mathcal{M}}^2)t}{1 - 2c(1 + \|(b, \omega)(T_0)\|_{\mathcal{M}}^2)t} \tag{3.40}$$

on $[T_0, T_0 + T_1]$ where

$$T_1 < \frac{1}{2c(1 + \|(b, \omega)(T_0)\|_{\mathcal{M}}^2)}.$$

Furthermore, (b^1, ω^1) must coincide with (b, ω) on $[T_0, T_0 + T_1] \cap [0, T]$ by uniqueness and the fact that they agree at $T_0 \in [0, T]$. Subsequently, we obtain from (3.40), strong right-continuity of $\|(b^1, \omega^1)\|_{\mathcal{M}}$ at time $t = T_0$ just as was done for $t = 0$ in Remark 3.2. Again, by uniqueness, this implies that $\|(b, \omega)\|_{\mathcal{M}}$ is also strongly right-continuous at time $t = T_0$. Since T_0 was chosen arbitrarily, this means that (b, ω) is strongly right-continuous on $[0, T]$. Since our system is reversible, we can repeat the above argument for the backward equation from which we obtain strong left-continuity on $[0, T]$. We have thus shown that weakly continuous solution (3.5) is indeed strongly continuous. □

3.4 Maximal Solution

We now end the section with a proof of the existence of a *maximal solution* to the TQG model.

Proof of Theorem 2.14 By Proposition 3.1, we found a time

$$T < \frac{1}{2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)} \tag{3.41}$$

with $c = c(\|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2})$ such that (b, ω, T) is a unique solution of (1.1)–(1.5). At time $T > 0$ given above, we now choose $(b_T, \omega_T) \in \mathcal{M}$, $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ and $f \in W^{2,2}(\mathbb{T}^2)$ as our new data where

$$b_T := b(T, \cdot), \quad \omega_T := \omega(T, \cdot).$$

By repeating the argument, we can also find

$$\tilde{T} < \frac{1}{2c(1 + \|(b_T, \omega_T)\|_{\mathcal{M}}^2)} \tag{3.42}$$

with $c = c(\|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2})$, such that with the new initial conditions (b_T, ω_T) , (b, ω) uniquely solves (1.1)–(1.5), in the sense of Definition 2.5, on $[T, T + \tilde{T}]$. By a gluing argument, we obtain a solution (b, ω, T_1) of (1.1)–(1.5) with $T_1 = T + \tilde{T}$. By this iterative procedure:

- we obtain an increasing family $(T_n)_{n \in \mathbb{N}}$ of time steps whose limit is $T_{\text{max}} \in (0, \infty]$;
- for each $n \in \mathbb{N}$, (b, ω, T_n) is a solution of (1.1)–(1.5);
- if $T_{\text{max}} < \infty$, then

$$\limsup_{T_n \rightarrow T_{\text{max}}} \|(b, \omega)(T_n)\|_{\mathcal{M}}^2 = \infty$$

since otherwise, we can repeat the procedure above to obtain $\tilde{\tilde{T}} > 0$ satisfying

$$\tilde{\tilde{T}} < \frac{1}{2c(1 + \|(b, \omega)(T_{\text{max}})\|_{\mathcal{M}}^2)}$$

with $c = c(\|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2})$, such that $(b, \omega, T_{\text{max}})$ is a solution of (1.1)–(1.5) with $T_{\text{max}}^{\text{max}} = T_{\text{max}} + \tilde{\tilde{T}}$. This will contradict the fact that T_{max} is the maximal time. \square

4 Convergence of α -TQG to TQG

In this section, for a given $\mathbf{u}_h \in W_{\text{div}}^{3,2}(\mathbb{T}^2; \mathbb{R}^2)$ and $f \in W^{2,2}(\mathbb{T}^2)$ serving as data for both the TQG (1.1)–(1.5) and the α -TQG (2.12)–(2.14), we aim to show that any family $(b^\alpha, \omega^\alpha, T_{\text{max}}^\alpha)_{\alpha > 0}$ of unique maximal solutions to the α -TQG model (2.12)–(2.14) converges strongly to the unique maximal solution $(b, \omega, T_{\text{max}})$ of the TQG model (1.1)–(1.5) as $\alpha \rightarrow 0$ provided that both system share the same initial conditions and are defined on a common time interval $[0, T]$. Obviously, the time horizon T should be smaller than T_{max}^α for any choice of α and smaller than T_{max} . To justify the existence of this common time horizon observe that Proposition 3.1 justifies the existence of a local solution on an interval $[0, T]$ for the TQG equation provided (3.4) holds, i.e.

that

$$T < \frac{1}{2c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2)}, \quad c = c(\|\mathbf{u}_h\|_{3,2}, \|f\|_{2,2});$$

For example, we can choose $T = 1/(4c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2))$. It follows that $T_{\max} > 1/(4c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2))$. The same argument applies to the α -TQG equation (see Lemma 2.17), that is, we also have $T_{\max}^\alpha > 1/(4c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2))$ for any value α . The common time horizon T can therefore be chosen to be $1/(4c(1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2))$. Our main result is the following.

Proposition 4.1 *Under Assumption 2.4, let $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)_{\alpha>0}$ be a family of unique maximal solutions of (2.12)–(2.14) and let (b, ω, T_{\max}) be the unique maximal solution of (1.1)–(1.5). Then*

$$\sup_{t \in [0, T]} (\|(b^\alpha - b)(t)\|_{1,2}^2 + \|(\omega^\alpha - \omega)(t)\|_2^2) \lesssim \alpha^2 T \exp(cT) [1 + \exp(cT)] \tag{4.1}$$

for any $T < \min\{T_{\max}, T_{\max}^\alpha\}$ and

$$\sup_{t \in [0, T]} \|(b^\alpha - b)(t)\|_{1,2} \rightarrow 0, \quad \sup_{t \in [0, T]} \|(\omega^\alpha - \omega)(t)\|_2 \rightarrow 0 \tag{4.2}$$

as $\alpha \rightarrow 0$.

We will now devote the entirety of this section to the proof of Proposition 4.1. In order to achieve this goal, we first introduce a unique maximal solution $(\bar{b}^\alpha, \bar{\omega}^\alpha, \bar{T}_{\max}^\alpha)$ with initial conditions $b_0 \in W^{3,2}(\mathbb{T}^2)$ and $\omega_0 \in W^{2,2}(\mathbb{T}^2)$ of the following ‘‘intermediate α -TQG’’ equation given by

$$\frac{\partial}{\partial t} \bar{b}^\alpha + \bar{\mathbf{u}}^\alpha \cdot \nabla \bar{b}^\alpha = 0, \tag{4.3}$$

$$\frac{\partial}{\partial t} \bar{\omega}^\alpha + \bar{\mathbf{u}}^\alpha \cdot \nabla (\bar{\omega}^\alpha - \bar{b}^\alpha) = -\mathbf{u}_h \cdot \nabla \bar{b}^\alpha, \tag{4.4}$$

where like (2.15),

$$\bar{\mathbf{u}}^\alpha = \nabla^\perp (1 - \alpha \Delta)^{-1} (\Delta - 1)^{-1} (\omega - f), \quad \mathbf{u}_h = \frac{1}{2} \nabla^\perp h, \tag{4.5}$$

and where ω (and not ω^α) satisfies the original potential vorticity equation (1.3), (1.4), where ω satisfies the elliptic equation in TQG. This means that

$$\bar{\mathbf{u}}^\alpha = \nabla^\perp (1 - \alpha \Delta)^{-1} (\Delta - 1)^{-1} (\omega - f) = \nabla^\perp (1 - \alpha \Delta)^{-1} \psi$$

with

$$\mathbf{u} = \nabla^\perp \psi$$

Then we obtain for any $k \geq 0$,

$$\begin{aligned} \|\bar{\mathbf{u}}^\alpha\|_{k,2}^2 &= \sum_{\xi \in \mathbb{Z}^2} \frac{(1 + |\xi|^2)^k |\xi|^2}{(1 + \alpha|\xi|^2)^2} |\hat{\psi}|^2 = \sum_{\xi \in \mathbb{Z}^2} \frac{(1 + |\xi|^2)^k}{(1 + \alpha|\xi|^2)^2} |\hat{\mathbf{u}}|^2 \\ &\leq \sum_{\xi \in \mathbb{Z}^2} (1 + |\xi|^2)^k |\hat{\mathbf{u}}|^2 = \|\mathbf{u}\|_{k,2}^2 \end{aligned}$$

So for solution \mathbf{u} in hand, we obtain a solution $\bar{\mathbf{u}}^\alpha$. The existence of a unique maximal solution to the above intermediate system, denoted by $\mathbf{i}\alpha$ -TQG for short, follows from the proof of existence for the original TQG and its particular, for a uniform-in- α set of data $(\mathbf{u}_h, f, b_0, \omega_0)$, we have the following uniform-in- α estimates

$$\begin{aligned} \sup_{\alpha > 0} \sup_{t \in [0, \bar{T}_{\max}^\alpha)} \|\bar{b}^\alpha\|_{3,2} &\lesssim 1, \\ \sup_{\alpha > 0} \sup_{t \in [0, \bar{T}_{\max}^\alpha)} \|\bar{\omega}^\alpha\|_{2,2} &\lesssim 1 \end{aligned} \tag{4.6}$$

for the maximal solution of (4.3)–(4.5). Indeed, exactly as in Theorem 2.15, we also have the following result.

Lemma 4.2 *Fix $\alpha > 0$. There exist a unique maximal solution $(\bar{b}^\alpha, \bar{\omega}^\alpha, \bar{T}_{\max}^\alpha)$ of (4.3)–(4.5) under Assumption 2.4. In particular, uniformly in $\alpha > 0$, the inequality*

$$\|(\bar{b}^\alpha, \bar{\omega}^\alpha)(t)\|_{\mathcal{M}} \lesssim 1 \tag{4.7}$$

holds for $t < \bar{T}_{\max}^\alpha$.

Remark 4.3 Since $1 \leq (1 + \alpha|\xi|^2)^2$ holds independently of $\alpha > 0$, it follows that

$$\begin{aligned} \|\bar{\mathbf{u}}^\alpha\|_{k+1,2}^2 &= \sum_{\xi \in \mathbb{Z}^2} \frac{(1 + |\xi|^2)^{k+1} |\xi|^2}{(1 + |\xi|^2)^2 (1 + \alpha|\xi|^2)^2} |\hat{\omega} - \hat{f}|^2 \\ &\leq \sum_{\xi \in \mathbb{Z}^2} \frac{(1 + |\xi|^2)^{k+1} |\xi|^2}{(1 + |\xi|^2)^2} |\hat{\omega} - \hat{f}|^2 \\ &\leq \sum_{\xi \in \mathbb{Z}^2} (1 + |\xi|^2)^k |\hat{\omega} - \hat{f}|^2 = \|\omega - f\|_{k,2}^2 \end{aligned}$$

Therefore, Lemma 2.2 also holds for the $\mathbf{i}\alpha$ -TQG model uniformly in α . Thus, the same analysis for the ‘normal’ TQG holds verbatim for the alpha TQG since the only place we find alpha is within the elliptic equation linking $\bar{\mathbf{u}}^\alpha$ with ω .

As a next step, we show that any family $(\bar{b}^\alpha, \bar{\omega}^\alpha, \bar{T}_{\max}^\alpha)_{\alpha > 0}$ of maximal solutions to the intermediate $\mathbf{i}\alpha$ -TQG model (4.3)–(4.5) (rather than of the α -TQG model (2.12)–(2.14)) converges strongly to the unique maximal solution of the TQG model (1.1)–(1.5) as $\alpha \rightarrow 0$ on the time interval $[0, T]$ where $T < \min\{\bar{T}_{\max}^\alpha, T_{\max}\}$. In order

to achieve this goal, we replicate the uniqueness argument in Sect. 3.3 by setting $b^{12} := \bar{b}^\alpha - b$, $\omega^{12} := \bar{\omega}^\alpha - \omega$ and $\mathbf{u}^{12} := \bar{\mathbf{u}}^\alpha - \mathbf{u}$ so that $(b^{12}, \mathbf{u}^{12}, \omega^{12})$ satisfies

$$\frac{\partial}{\partial t} b^{12} + \mathbf{u}^{12} \cdot \nabla \bar{b}^\alpha + \mathbf{u} \cdot \nabla b^{12} = 0, \tag{4.8}$$

$$\frac{\partial}{\partial t} \omega^{12} + \mathbf{u}^{12} \cdot \nabla (\bar{\omega}^\alpha - \bar{b}^\alpha) + \mathbf{u} \cdot \nabla (\omega^{12} - b^{12}) = -\mathbf{u}_h \cdot \nabla b^{12}. \tag{4.9}$$

Similar to the uniqueness argument, if we apply ∂^β to the equation for b^{12} above where now $|\beta| \leq 1$, we obtain

$$\frac{\partial}{\partial t} \partial^\beta b^{12} + \mathbf{u} \cdot \nabla \partial^\beta b^{12} = S_1 + S_2 - \mathbf{u}^{12} \cdot \nabla \partial^\beta \bar{b}^\alpha, \tag{4.10}$$

where

$$\begin{aligned} S_1 &:= \mathbf{u}^{12} \cdot \partial^\beta \nabla \bar{b}^\alpha - \partial^\beta (\mathbf{u}^{12} \cdot \nabla \bar{b}^\alpha) = -\partial \mathbf{u}^{12} \cdot \nabla \bar{b}^\alpha, \\ S_2 &:= \mathbf{u} \cdot \partial^\beta \nabla b^{12} - \partial^\beta (\mathbf{u} \cdot \nabla b^{12}) = -\partial \mathbf{u} \cdot \nabla b^{12} \end{aligned}$$

are such that

$$\|S_1\|_2 \lesssim \|\mathbf{u}^{12}\|_{1,2} \|\bar{b}^\alpha\|_{3,2}, \quad \|S_2\|_2 \lesssim \|\mathbf{u}\|_{3,2} \|b^{12}\|_{1,2}. \tag{4.11}$$

Also,

$$\left| \langle \mathbf{u}^{12} \cdot \nabla \partial^\beta \bar{b}^\alpha, \partial^\beta b^{12} \rangle \right| \lesssim \|\mathbf{u}^{12}\|_{1,2} \|\bar{b}^\alpha\|_{3,2} \|b^{12}\|_{1,2}. \tag{4.12}$$

Collecting the information above yields the estimate

$$\frac{d}{dt} \|b^{12}\|_{1,2}^2 \lesssim (\|\mathbf{u}^{12}\|_{1,2} \|\bar{b}^\alpha\|_{3,2} + \|\mathbf{u}\|_{3,2} \|b^{12}\|_{1,2}) \|b^{12}\|_{1,2}. \tag{4.13}$$

For the equation for ω^{12} , we aim to derive a square-integrable estimate in space. For this, we first note that

$$\left| \langle \mathbf{u} \cdot \nabla \omega^{12}, \omega^{12} \rangle \right| = 0. \tag{4.14}$$

Next, we have that

$$\left| \langle (\mathbf{u} - \mathbf{u}_h) \cdot \nabla b^{12}, \omega^{12} \rangle \right| \lesssim \|\mathbf{u} - \mathbf{u}_h\|_\infty \|b^{12}\|_{1,2} \|\omega^{12}\|_2. \tag{4.15}$$

Finally, we also have that

$$\begin{aligned} \left| \langle \mathbf{u}^{12} \cdot \nabla (\bar{\omega}^\alpha - \bar{b}^\alpha), \omega^{12} \rangle \right| &\lesssim \|\mathbf{u}^{12}\|_4 \|\nabla (\bar{\omega}^\alpha - \bar{b}^\alpha)\|_4 \|\omega^{12}\|_2 \\ &\lesssim \|\mathbf{u}^{12}\|_{1,2} \|\bar{\omega}^\alpha - \bar{b}^\alpha\|_{2,2} \|\omega^{12}\|_2. \end{aligned} \tag{4.16}$$

Therefore,

$$\frac{d}{dt} \|\omega^{12}\|_2^2 \lesssim (\|\mathbf{u} - \mathbf{u}_h\|_\infty \|b^{12}\|_{1,2} + \|\mathbf{u}^{12}\|_{1,2} \|\bar{\omega}^\alpha - \bar{b}^\alpha\|_{2,2}) \|\omega^{12}\|_2. \tag{4.17}$$

Summing up with the estimate for b^{12} then yields

$$\begin{aligned} \frac{d}{dt} (\|b^{12}\|_{1,2}^2 + \|\omega^{12}\|_2^2) &\lesssim (\|\mathbf{u}^{12}\|_{1,2} \|\bar{b}^\alpha\|_{3,2} + \|\mathbf{u}\|_{3,2} \|b^{12}\|_{1,2}) \|b^{12}\|_{1,2} \\ &\quad + (\|\mathbf{u} - \mathbf{u}_h\|_\infty \|b^{12}\|_{1,2} + \|\mathbf{u}^{12}\|_{1,2} \|\bar{\omega}^\alpha - \bar{b}^\alpha\|_{2,2}) \|\omega^{12}\|_2. \end{aligned} \tag{4.18}$$

If we use (2.7) and (4.7), we obtain

$$\begin{aligned} \frac{d}{dt} (\|b^{12}\|_{1,2}^2 + \|\omega^{12}\|_2^2) &\lesssim (\|\mathbf{u}^{12}\|_{1,2} + \|b^{12}\|_{1,2}) \|b^{12}\|_{1,2} \\ &\quad + (\|b^{12}\|_{1,2} + \|\mathbf{u}^{12}\|_{1,2}) \|\omega^{12}\|_2 \\ &\lesssim \|\mathbf{u}^{12}\|_{1,2}^2 + \|b^{12}\|_{1,2}^2 + \|\omega^{12}\|_2^2. \end{aligned} \tag{4.19}$$

Now recall that the velocity fields for the α -TQG, the $\mathbf{i}\alpha$ -TQG, and the TQG are given by

$$\mathbf{u}^\alpha = \nabla^\perp (1 - \alpha \Delta)^{-1} (\Delta - 1)^{-1} (\omega^\alpha - f), \tag{4.20}$$

$$\bar{\mathbf{u}}^\alpha = \nabla^\perp (1 - \alpha \Delta)^{-1} (\Delta - 1)^{-1} (\omega - f), \tag{4.21}$$

$$\mathbf{u} = \nabla^\perp (\Delta - 1)^{-1} (\omega - f), \tag{4.22}$$

respectively. As a result, in particular, the difference

$$\bar{\mathbf{u}}^\alpha - \mathbf{u} = \alpha \Delta \nabla^\perp (1 - \alpha \Delta)^{-1} (\Delta - 1)^{-1} (\omega - f)$$

enjoys two extra order of regularity. Furthermore, since each of these individual velocity fields satisfies the bound (2.7), it immediately follows that

$$\|\bar{\mathbf{u}}^\alpha - \mathbf{u}\|_{1,2}^2 \lesssim \alpha^2 \|\mathbf{u}\|_{3,2}^2 \lesssim \alpha^2 \|\omega - f\|_{2,2}^2 \lesssim \alpha^2 (\|\omega\|_{2,2}^2 + \|f\|_{2,2}^2) \lesssim \alpha^2. \tag{4.23}$$

This means that for the comparison of the TQG and $\mathbf{i}\alpha$ -TQG, we can conclude from (4.19) that

$$\frac{d}{dt} (\|\bar{b}^\alpha - b\|_{1,2}^2 + \|\bar{\omega}^\alpha - \omega\|_2^2) \lesssim \alpha^2 + \|\bar{b}^\alpha - b\|_{1,2}^2 + \|\bar{\omega}^\alpha - \omega\|_2^2 \tag{4.24}$$

and by Grönwall’s lemma, (note that $b_0^{12} = 0$ and $\omega_0^{12} = 0$)

$$\|(\bar{b}^\alpha - b)(t)\|_{1,2}^2 + \|(\bar{\omega}^\alpha - \omega)(t)\|_2^2 \lesssim \alpha^2 T \exp(cT), \quad t \in [0, T]. \tag{4.25}$$

The above gives the decay rate of the difference of the solution to the $\mathbf{i}\alpha$ -TQG and TQG equations. Our next goal is to obtain a decay rate for the difference between the $\mathbf{i}\alpha$ -TQG and the α -TQG equations. Recall that they share the same initial data. For this, we use the estimate

$$\|\bar{\mathbf{u}}^\alpha - \mathbf{u}^\alpha\|_{1,2}^2 \lesssim \|\omega - \omega^\alpha\|_2^2 \lesssim \|\bar{\omega}^\alpha - \omega\|_2^2 + \|\bar{\omega}^\alpha - \omega^\alpha\|_2^2$$

so that by setting $b^{12} := \bar{b}^\alpha - b^\alpha$, $\omega^{12} := \bar{\omega}^\alpha - \omega^\alpha$ and $\mathbf{u}^{12} := \bar{\mathbf{u}}^\alpha - \mathbf{u}^\alpha$, we obtain from (4.19),

$$\begin{aligned} \frac{d}{dt} (\|b^{12}\|_{1,2}^2 + \|\omega^{12}\|_2^2) &\lesssim \|\mathbf{u}^{12}\|_{1,2}^2 + \|b^{12}\|_{1,2}^2 + \|\omega^{12}\|_2^2 \\ &\lesssim \|\bar{\omega}^\alpha - \omega\|_2^2 + \|b^{12}\|_{1,2}^2 + \|\omega^{12}\|_2^2. \end{aligned} \tag{4.26}$$

Since both systems share the same initial conditions, by Grönwall’s lemma and (4.25), it follows that

$$\begin{aligned} \|(\bar{b}^\alpha - b^\alpha)(t)\|_{1,2}^2 + \|(\bar{\omega}^\alpha - \omega^\alpha)(t)\|_2^2 &\leq \exp(cT) \int_0^T c \|(\bar{\omega}^\alpha - \omega)(t)\|_2^2 dt \\ &\lesssim \alpha^2 T [\exp(cT)]^2. \end{aligned} \tag{4.27}$$

By the triangle inequality, it follows from (4.25) and (4.27) that

$$\|(b^\alpha - b)(t)\|_{1,2}^2 + \|(\omega^\alpha - \omega)(t)\|_2^2 \lesssim \alpha^2 T \exp(cT) [1 + \exp(cT)] \tag{4.28}$$

so that

$$\sup_{t \in [0, T]} \|(b^\alpha - b)(t)\|_{1,2} \rightarrow 0, \quad \sup_{t \in [0, T]} \|(\omega^\alpha - \omega)(t)\|_2 \rightarrow 0$$

as $\alpha \rightarrow 0$. This ends the proof.

Remark 4.4 We conjecture that Proposition 4.1 may be extended to the stronger space $W^{2,2}(\mathbb{T}^2) \times W^{1,2}(\mathbb{T}^2)$ in which we showed the stability and uniqueness of the maximal solution. The cost, however, may be the loss of the corresponding decay rate (4.1). The ideas in the proof of Proposition 6 of Terence Tao’s note: <https://terrytao.wordpress.com/2018/10/09/254a-notes-3-local-well-posedness-for-the-euler-equations/#more-10770> may suffice.

5 Conditions for Blow-up of the α -TQG Model

In the following, we give Beale et al. (1984) (BKM)-type conditions under which we expect the strong solution of the α -TQG to blow-up.

Theorem 5.1 Fix $\alpha > 0$. Under Assumption 2.4, let $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)$ be a maximal solution of (2.12)–(2.14). If $T_{\max}^\alpha < \infty$, then

$$\int_0^{T_{\max}^\alpha} \|\nabla b^\alpha\|_\infty dt = \infty.$$

Proof First of all, fix $\alpha > 0$ and let $T_{\max}^\alpha > 0$ be the maximal time so that

$$\limsup_{T_n \rightarrow T_{\max}^\alpha} \|(b^\alpha, \omega^\alpha)(T_n)\|_{\mathcal{M}}^2 = \infty. \tag{5.1}$$

We now suppose that

$$\int_0^{T_{\max}^\alpha} \|\nabla b^\alpha\|_\infty dt = K < \infty \tag{5.2}$$

and show that

$$\|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2 \lesssim 1, \quad t < T_{\max}^\alpha \tag{5.3}$$

holds yielding a contradiction to (5.1).

Before we show the estimate (5.3), we first require preliminary estimates for $\|\nabla \mathbf{u}^\alpha\|_\infty$ and $\|\omega^\alpha\|_2$. To obtain these estimates, we note that the Fourier multiplier $m^\alpha(\xi)$ defined below satisfies the bound

$$m^\alpha(\xi) := \frac{(1 + |\xi|^2)|\xi|^2}{(1 + \alpha|\xi|^2)^2} \lesssim_\alpha 1 \tag{5.4}$$

for all fixed $\alpha > 0$ and in particular, the bound may be taken uniformly of all $\alpha \geq 1$. Due to (5.4), the continuous embedding $W^{3,2}(\mathbb{T}^2) \hookrightarrow W^{1,\infty}(\mathbb{T}^2)$ and Plancherel’s identity, we can conclude that

$$\|\nabla \mathbf{u}^\alpha\|_\infty \lesssim \|\mathbf{u}^\alpha\|_{3,2} \lesssim_\alpha \|\omega^\alpha - f\|_2 \lesssim \|\omega^\alpha - f\|_{2,2}. \tag{5.5}$$

On the other hand, if we test (2.13) with ω^α and use (2.7) for $w = (1 - \alpha \Delta)^{-1}(\omega^\alpha - f)$ and $k = 0$, we obtain

$$\frac{d}{dt} \|\omega^\alpha\|_2^2 \lesssim \left(\|\mathbf{u}^\alpha\|_2 \|\omega^\alpha\|_2 + \|\mathbf{u}_h\|_2 \|\omega^\alpha\|_2 \right) \|\nabla b^\alpha\|_\infty \lesssim \left(1 + \|\omega^\alpha\|_2^2 \right) \|\nabla b^\alpha\|_\infty \tag{5.6}$$

for a constant depending only on $\|\mathbf{u}_h\|_2$ and $\|f\|_2$. It therefore follows from (5.6) and (5.2) that

$$\|\omega^\alpha(t)\|_2^2 \lesssim (1 + \|\omega_0\|_2^2) \exp(cK), \quad t < T_{\max}^\alpha \tag{5.7}$$

which when combined with the first estimate in (5.5) yields

$$\|\nabla \mathbf{u}^\alpha(t)\|_\infty \lesssim_\alpha 1 + (1 + \|\omega_0\|_2^2) \exp(cK) \lesssim_\alpha (1 + \|\omega_0\|_2^2) \exp(cK), \quad t < T_{\max}^\alpha. \tag{5.8}$$

Recall that $f \in W^{2,2}(\mathbb{T}^2)$ by assumption.

With these preliminary estimates in hand, we can proceed to derive estimates for $(b^\alpha, \omega^\alpha)$ in the space \mathcal{M} of existence of a solution. Since the space of smooth functions is dense in \mathcal{M} , it suffices to show our result for a smooth solution pair $(b^\alpha, \omega^\alpha)$.

To achieve our goal, we apply ∂^β to (2.12) for $|\beta| \leq 3$ to obtain

$$\frac{\partial}{\partial t} \partial^\beta b^\alpha + \mathbf{u}^\alpha \cdot \nabla \partial^\beta b^\alpha = R_1 \tag{5.9}$$

where

$$R_1 := \mathbf{u}^\alpha \cdot \partial^\beta \nabla b^\alpha - \partial^\beta (\mathbf{u}^\alpha \cdot \nabla b^\alpha).$$

Now since $\operatorname{div} \mathbf{u}^\alpha = 0$, if we multiply (5.9) by $\partial^\beta b^\alpha$ and sum over the multiindex β so that $|\beta| \leq 3$, we obtain

$$\begin{aligned} \frac{d}{dt} \|b^\alpha\|_{3,2}^2 &\lesssim \left(\|\nabla \mathbf{u}^\alpha\|_\infty \|b^\alpha\|_{3,2} + \|\nabla b^\alpha\|_\infty \|\mathbf{u}^\alpha\|_{3,2} \right) \|b^\alpha\|_{3,2} \\ &\lesssim (\|\nabla \mathbf{u}^\alpha\|_\infty + \|\nabla b^\alpha\|_\infty) (1 + \|(b^\alpha, \omega^\alpha)\|_{\mathcal{M}}^2) \end{aligned} \tag{5.10}$$

where we have used (2.7) for $w = (1 - \alpha \Delta)^{-1}(\omega^\alpha - f)$ and $k = 2$.

Next, we find a bound for $\|\omega^\alpha\|_{2,2}^2$. For this, we apply ∂^β to (2.13) for $|\beta| \leq 2$ and we obtain

$$\frac{\partial}{\partial t} \partial^\beta \omega^\alpha + \mathbf{u}^\alpha \cdot \nabla \partial^\beta (\omega^\alpha - b^\alpha) + \mathbf{u}_h \cdot \nabla \partial^\beta b^\alpha = R_2 + R_3 + R_4 \tag{5.11}$$

where

$$\begin{aligned} R_2 &:= \mathbf{u}^\alpha \cdot \partial^\beta \nabla \omega^\alpha - \partial^\beta (\mathbf{u}^\alpha \cdot \nabla \omega^\alpha), \\ R_3 &:= -\mathbf{u}^\alpha \cdot \partial^\beta \nabla b^\alpha + \partial^\beta (\mathbf{u}^\alpha \cdot \nabla b^\alpha), \\ R_4 &:= \mathbf{u}_h \cdot \partial^\beta \nabla b^\alpha - \partial^\beta (\mathbf{u}_h \cdot \nabla b^\alpha) \end{aligned}$$

are such that by using Lemma 2.3, Sobolev embeddings, the second inequality in (5.5) (to obtain the term $1 + \|\omega^\alpha\|_2$ in R_2 below) and (2.7), the estimates

$$\begin{aligned} |\langle R_2, \partial^\beta \omega^\alpha \rangle| &\lesssim \|R_2\|_2 \|\omega^\alpha\|_{2,2} \\ &\lesssim \left(\|\nabla \mathbf{u}^\alpha\|_\infty \|\omega^\alpha\|_{2,2} + \|\omega^\alpha\|_4 \|\mathbf{u}^\alpha\|_{2,4} \right) \|\omega^\alpha\|_{2,2} \\ &\lesssim \left(\|\nabla \mathbf{u}^\alpha\|_\infty \|\omega^\alpha\|_{2,2} + \|\omega^\alpha\|_{2,2} \|\mathbf{u}^\alpha\|_{3,2} \right) \|\omega^\alpha\|_{2,2} \end{aligned}$$

$$\lesssim \left(\|\nabla \mathbf{u}^\alpha\|_\infty \|\omega^\alpha\|_{2,2} + \|\omega^\alpha\|_{2,2} (1 + \|\omega^\alpha\|_2) \right) \|\omega^\alpha\|_{2,2}, \tag{5.12}$$

$$\begin{aligned} |\langle R_3, \partial^\beta \omega^\alpha \rangle| &\lesssim \|R_3\|_2 \|\omega^\alpha\|_{2,2} \\ &\lesssim \left(\|\nabla \mathbf{u}^\alpha\|_\infty \|b^\alpha\|_{3,2} + \|\nabla b^\alpha\|_\infty \|\mathbf{u}^\alpha\|_{2,2} \right) \|\omega^\alpha\|_{2,2} \\ &\lesssim \left(\|\nabla \mathbf{u}^\alpha\|_\infty \|b^\alpha\|_{3,2} + \|\nabla b^\alpha\|_\infty (1 + \|\omega^\alpha\|_{2,2}) \right) \|\omega^\alpha\|_{2,2}, \end{aligned} \tag{5.13}$$

$$\begin{aligned} |\langle R_4, \partial^\beta \omega^\alpha \rangle| &\lesssim \|R_4\|_2 \|\omega^\alpha\|_{2,2} \\ &\lesssim \left(\|\nabla \mathbf{u}_h\|_\infty \|b^\alpha\|_{3,2} + \|\nabla b^\alpha\|_\infty \|\mathbf{u}_h\|_{2,2} \right) \|\omega^\alpha\|_{2,2} \\ &\lesssim \left(\|b^\alpha\|_{3,2} + \|\nabla b^\alpha\|_\infty \right) \|\omega^\alpha\|_{2,2} \end{aligned} \tag{5.14}$$

hold true. Recall that $\mathbf{u}_h \in W^{3,2}(\mathbb{T}^2)$ by assumption. Next, by using $\operatorname{div} \mathbf{u}^\alpha = 0$, we have the following identity

$$\langle (\mathbf{u}^\alpha \cdot \nabla \partial^\beta \omega^\alpha), \partial^\beta \omega^\alpha \rangle = \frac{1}{2} \int_{\mathbb{T}^2} \operatorname{div}(\mathbf{u}^\alpha |\partial^\beta \omega^\alpha|^2) \, dx = 0. \tag{5.15}$$

Additionally, the following estimates for L^2 inner products hold true

$$\left| \langle (\mathbf{u}^\alpha \cdot \nabla \partial^\beta b^\alpha), \partial^\beta \omega^\alpha \rangle \right| \lesssim \|\mathbf{u}^\alpha\|_\infty \|b^\alpha\|_{3,2}^2 + \|\mathbf{u}^\alpha\|_\infty \|\omega^\alpha\|_{2,2}^2, \tag{5.16}$$

$$\left| \langle (\mathbf{u}_h \cdot \nabla \partial^\beta b^\alpha), \partial^\beta \omega^\alpha \rangle \right| \lesssim \|b^\alpha\|_{3,2}^2 + \|\omega^\alpha\|_{2,2}^2. \tag{5.17}$$

since $\mathbf{u}_h \in W^{3,2}(\mathbb{T}^2)$. If we now collect the estimates above (keeping in mind that $f \in W^{2,2}(\mathbb{T}^2)$ and $\mathbf{u}_h \in W^{3,2}(\mathbb{T}^2)$), we obtain by multiplying (5.11) by $\partial^\beta \omega^\alpha$ with $|\beta| \leq 2$ and then summing over $|\beta| \leq 2$, the following

$$\frac{d}{dt} \|\omega^\alpha\|_{2,2}^2 \lesssim (1 + \|\mathbf{u}^\alpha\|_{1,\infty} + \|\nabla b^\alpha\|_\infty + \|\omega^\alpha\|_2) (1 + \|(b^\alpha, \omega^\alpha)\|_{\mathcal{M}}^2). \tag{5.18}$$

Summing up (5.10) and (5.18) and using (5.7)–(5.5) yields

$$\frac{d}{dt} (1 + \|(b^\alpha, \omega^\alpha)\|_{\mathcal{M}}^2) \lesssim_\alpha \left((1 + \|\omega_0\|_2^2) \exp(cK) + \|\nabla b^\alpha\|_\infty \right) (1 + \|(b^\alpha, \omega^\alpha)\|_{\mathcal{M}}^2) \tag{5.19}$$

so that by Grönwall’s lemma and (5.2), we obtain

$$\begin{aligned} \|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2 &\leq [1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2] \exp \left(c(\alpha) T_{\max} (1 + \|\omega_0\|_2^2) \exp(cK) + c(\alpha) K \right) \\ &\lesssim_{T_{\max}, K, f, \mathbf{u}_h, \omega_0, b_0, \alpha} 1 \end{aligned} \tag{5.20}$$

for all $t < T_{\max}^\alpha$ contradicting (5.1). □

Remark 5.2 (Global existence for constant-in-space buoyancy) We note that for the α -TQG model, when the buoyancy is constant in space so that $\nabla b^\alpha = 0$, then the

equation for the buoyancy decouples from that of the potential vorticity. The potential vorticity equation then reduces to the two-dimensional Euler equation albeit a different Biot–Savart law. Nevertheless, by inspecting the proof of Theorem 5.1 above, one observes that the same global-in-time result applies. In particular, the residual estimate (5.12) still holds true and we obtain from this, a refine version of (5.7)–(5.8) where $\exp(cK) = 1$. Recall from (5.2) that $K = 0$ when $\nabla b^\alpha = 0$.

Remark 5.3 (Global existence for super diffusive buoyancy) Fix $\alpha > 0$. Note that by adding any super diffusive term $(-\Delta)^{\beta/2} b^\alpha$, $\beta > 2$ to the right-hand side of (2.12) so that

$$\int_0^t \|\nabla b^\alpha(s)\|_\infty ds \lesssim \int_0^t \|(-\Delta)^{\beta/2} b^\alpha(s)\|_2^2 ds$$

holds for all $t < T_{\max}^\alpha$, then we obtain a global solution since in this case, we obtain the energy estimate

$$\|b^\alpha(t)\|_2^2 + \int_0^t \|(-\Delta)^{\beta/2} b^\alpha(s)\|_2^2 ds \leq \|b_0\|_2 \leq \|b_0\|_{3,2}$$

for all $t < T_{\max}^\alpha$.

Next, we also give an alternative to the blow-up condition in Theorem 5.1 above.

Theorem 5.4 Fix $\alpha > 0$. Under Assumption 2.4, let $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)$ be a maximal solution of (2.12)–(2.14). If $T_{\max}^\alpha < \infty$, then

$$\int_0^{T_{\max}^\alpha} \|\nabla \mathbf{u}^\alpha\|_\infty dt = \infty.$$

Proof In analogy with the proof of Theorem 5.1, we fix $\alpha > 0$, let $T_{\max}^\alpha > 0$ be the maximal time so that

$$\limsup_{T_n \rightarrow T_{\max}^\alpha} \|(b^\alpha, \omega^\alpha)(T_n)\|_{\mathcal{M}}^2 = \infty. \tag{5.21}$$

We now suppose that

$$\int_0^{T_{\max}^\alpha} \|\nabla \mathbf{u}^\alpha\|_\infty dt = K < \infty \tag{5.22}$$

and show that

$$\|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2 \lesssim 1, \quad t < T_{\max}^\alpha \tag{5.23}$$

holds, yielding a contradiction to (5.21).

To show this, we first fix $\alpha > 0$. If we differentiate (2.12) in space, multiply the resulting equation by $p|\nabla b^\alpha|^{p-2}\nabla b^\alpha$ where $p > 1$ is fixed and finite in this instant, and then integrate over \mathbb{T}^2 , we obtain

$$\frac{d}{dt} \|\nabla b^\alpha\|_p^p \leq p \|\nabla \mathbf{u}^\alpha\|_\infty \|\nabla b^\alpha\|_p^p. \tag{5.24}$$

Rather than use Grönwall’s lemma at this point, we use the analogous separation of variable technique to obtain

$$\frac{1}{p} \ln \|\nabla b^\alpha(t)\|_p^p \leq \frac{1}{p} \ln \|\nabla b_0\|_p^p + \int_0^t \|\nabla \mathbf{u}^\alpha(s)\|_\infty ds$$

so that

$$\|\nabla b^\alpha(t)\|_p \leq \|\nabla b_0\|_p \exp\left(\int_0^t \|\nabla \mathbf{u}^\alpha(s)\|_\infty ds\right)$$

holds for all $t < T_{\max}^\alpha$ uniformly in $p > 1$. Since $b_0 \in W^{3,2}(\mathbb{T}^2)$, by Sobolev embedding and (5.22), we obtain

$$\|\nabla b^\alpha(t)\|_\infty \lesssim \|b_0\|_{3,2} \exp(K) \tag{5.25}$$

for all $t < T_{\max}^\alpha$ so that

$$\int_0^{T_{\max}^\alpha} \|\nabla b^\alpha(t)\|_\infty dt \leq c T_{\max}^\alpha \|b_0\|_{3,2} \exp(K).$$

We can therefore deduce from the estimate (5.6) that

$$\|\omega^\alpha(t)\|_2^2 \lesssim (1 + \|\omega_0\|_2^2) \exp(c T_{\max}^\alpha \|b_0\|_{3,2} \exp(K)), \quad t < T_{\max}^\alpha. \tag{5.26}$$

Summing up (5.10) and (5.18) and using (5.25)–(5.26) yields

$$\frac{d}{dt} (1 + \|(b^\alpha, \omega^\alpha)\|_{\mathcal{M}}^2) \lesssim (1 + \|\mathbf{u}^\alpha\|_{1,\infty} + \mathfrak{A})(1 + \|(b^\alpha, \omega^\alpha)\|_{\mathcal{M}}^2) \tag{5.27}$$

where

$$\mathfrak{A} := \|b_0\|_{3,2} \exp(K) + (1 + \|\omega_0\|_2^2) \exp(c T_{\max}^\alpha \|b_0\|_{3,2} \exp(K)).$$

We can therefore conclude from Grönwall’s lemma and (5.22) that

$$\begin{aligned} \|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2 &\leq [1 + \|(b_0, \omega_0)\|_{\mathcal{M}}^2] \exp\left(c T_{\max}^\alpha + cK + c T_{\max}^\alpha \mathfrak{A}\right) \\ &\lesssim_{T_{\max}^\alpha, K, f, \mathbf{u}_h, \omega_0, b_0} 1 \end{aligned} \tag{5.28}$$

for all $t < T_{\max}^\alpha$ contradicting (5.21). □

In order to give the final blow-up criterion, let us now recall the following endpoint Sobolev inequality by Brezis and Gallouet (1979).

Lemma 5.5 *If $f \in W^{2,2}(\mathbb{T}^2)$ then*

$$\|f\|_\infty \lesssim (1 + \|f\|_{1,2})\sqrt{\ln(e + \|f\|_{2,2})}.$$

Remark 5.6 Note that the original statement in Brezis and Gallouet (1979) restricted the size of $\|f\|_{1,2}$ to being at most one. The current form for any size of $\|f\|_{1,2}$ follows immediately as demonstrated in, for example, (Wang and Zhang 2011; Dinvey 2020).

With Lemma 5.5 in hand, we can now show the final blow-up condition.

Theorem 5.7 *Fix $\alpha > 0$. Let $(b^\alpha, \omega^\alpha, T_{\max}^\alpha)$ be a maximal solution of (2.12)–(2.14). If $T_{\max}^\alpha < \infty$, then*

$$\int_0^{T_{\max}^\alpha} \|\nabla b^\alpha\|_{1,2} dt = \infty.$$

Proof Fix $\alpha > 0$ and let $T_{\max}^\alpha > 0$ be the maximal time so that

$$\limsup_{T_n \rightarrow T_{\max}^\alpha} \|(b^\alpha, \omega^\alpha)(T_n)\|_{\mathcal{M}}^2 = \infty. \tag{5.29}$$

We now suppose that

$$\int_0^{T_{\max}^\alpha} \|\nabla b^\alpha\|_{1,2} dt = K < \infty \tag{5.30}$$

and show that

$$\|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2 \lesssim 1, \quad t < T_{\max}^\alpha \tag{5.31}$$

holds and thereby yields a contradiction to (5.29).

To show this, we first fix $\alpha > 0$ and define

$$g(t) := e + \|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2, \quad t > 0.$$

With this definition in hand, it follows from estimates (5.10) and (5.18) that

$$\frac{d}{dt}g \lesssim (1 + \|\mathbf{u}^\alpha\|_{1,\infty} + \|\omega^\alpha\|_2 + \|\nabla b^\alpha\|_\infty)g. \tag{5.32}$$

By using Lemma 5.5 and the monotonic property of logarithms however, we can deduce that

$$\|\nabla b^\alpha\|_\infty \lesssim (1 + \|\nabla b^\alpha\|_{1,2}) \ln g. \tag{5.33}$$

Next, we observe that the second estimate in (5.5) yields

$$1 + \|\mathbf{u}^\alpha\|_{1,\infty} + \|\omega^\alpha\|_2 \lesssim_\alpha 1 + \|\omega^\alpha\|_2 \tag{5.34}$$

for $f \in L^2(\mathbb{T}^2)$. However, if we test (2.13) with ω^α , we also obtain

$$\frac{d}{dt} \|\omega^\alpha\|_2^2 \lesssim \left(\|\mathbf{u}^\alpha\|_\infty \|\omega^\alpha\|_2 + \|\mathbf{u}_h\|_\infty \|\omega^\alpha\|_2 \right) \|\nabla b^\alpha\|_2 \lesssim \left(1 + \|\omega^\alpha\|_2^2 \right) \|\nabla b^\alpha\|_{1,2} \tag{5.35}$$

for a constant depending only on $\|\mathbf{u}_h\|_\infty$ and $\|f\|_2$. It therefore follows from (5.30) that

$$\|\omega^\alpha(t)\|_2^2 \lesssim \left(1 + \|\omega_0\|_2^2 \right) \exp(cK), \quad t < T_{\max}^\alpha. \tag{5.36}$$

Using the fact that $\ln g \geq 1$, we can conclude from (5.34) and (5.36) that

$$1 + \|\mathbf{u}^\alpha\|_{1,\infty} + \|\omega^\alpha\|_2 \lesssim_\alpha \left(1 + \|\omega_0\|_2^2 \right) \exp(cK) \ln g. \tag{5.37}$$

If we now combine this estimate with (5.33), we can conclude from (5.32) that

$$\frac{d}{dt} g \lesssim \left(1 + \|\omega_0\|_2 + \|\nabla b^\alpha\|_{1,2} \right) \exp(cK) g \ln g$$

so that

$$g(t) \leq g(0) \exp \int_0^t c(1 + \|\omega_0\|_2 + \|\nabla b^\alpha(s)\|_{1,2}) \exp(cK) \, ds$$

holds for $t < T_{\max}^\alpha$. In particular, given (5.30), we have shown that

$$\begin{aligned} \|(b^\alpha, \omega^\alpha)(t)\|_{\mathcal{M}}^2 &\leq \left[e + \|(b_0^\alpha, \omega_0^\alpha)\|_{\mathcal{M}}^2 \right] \exp(c(T_{\max}^\alpha + K) \exp(cK)) \\ &\lesssim_{T_{\max}^\alpha, K, f, \mathbf{u}_h, \omega_0, b_0} 1 \end{aligned} \tag{5.38}$$

for all $t < T_{\max}^\alpha$ contradicting (5.29). □

6 Numerical Implementation

For numerical implementation, we choose to work in weaker spaces than what the well-posedness theorem dictates. Additionally, we allow for boundary conditions in the numerical setup. We recognise these choices create gaps between the theory and the implementation.

6.1 Discretisation Methods for α -TQG and TQG

In this subsection, we describe the finite element (FEM) spatial discretisation and finite difference Runge–Kutta time stepping discretisation methods that are utilised

for the α -TQG system. By setting α to zero, we obtain our numerical setup for the TQG system.

Consider a bounded domain \mathcal{D} . Let $\partial\mathcal{D}$ denote the boundary. We impose Dirichlet boundary conditions

$$\psi^\alpha = 0, \quad \Delta\psi^\alpha = 0, \quad \text{on } \partial\mathcal{D}. \tag{6.1}$$

For $\mathcal{D} = \mathbb{T}^2$, our numerical setup does not change, in which case the boundary flux terms in the discretised equations are set to zero.

6.1.1 The Stream Function Equation

Let $H^1(\mathcal{D})$ denote the Sobolev $W^{1,2}(\mathcal{D})$ space and let $\|\cdot\|_{\partial\mathcal{D}}$ denote the $L^2(\partial\mathcal{D})$ norm. Define the space

$$W^1(\mathcal{D}) := \left\{ v \in H^1(\mathcal{D}) \mid \|v\|_{\partial\mathcal{D}} = 0 \right\}. \tag{6.2}$$

We express (2.14) as two inhomogeneous Helmholtz equations

$$\omega^\alpha - f = (\Delta - 1)\tilde{\psi} \tag{6.3}$$

$$-\tilde{\psi} = (\alpha\Delta - 1)\psi^\alpha. \tag{6.4}$$

We take $\tilde{\psi}, \psi^\alpha \in W^1(\mathcal{D})$. Note that as α tends to zero, the Biot–Savart law for $(\omega^\alpha, \psi^\alpha)$ and the Biot–Savart law $(\omega^\alpha, \tilde{\psi})$ coincide. This removes the need for the additional boundary condition in (6.1). Since the two equations are of the same form, let us first consider (6.3). Using an arbitrary test function $\phi \in W^1(\mathcal{D})$, we obtain the following weak form of (6.3),

$$\langle \nabla\tilde{\psi}, \nabla\phi \rangle_{\mathcal{D}} + \langle \tilde{\psi}, \phi \rangle_{\mathcal{D}} = -\langle \omega^\alpha - f, \phi \rangle_{\mathcal{D}}. \tag{6.5}$$

Define the functionals

$$L_\alpha(v, \phi) := \alpha \langle \nabla v, \nabla\phi \rangle_{\mathcal{D}} + \langle v, \phi \rangle_{\mathcal{D}} \tag{6.6}$$

$$F_\alpha(\phi) := -\langle \cdot, \phi \rangle_{\mathcal{D}} \tag{6.7}$$

for $v, \phi \in W^1(\mathcal{D})$, then (6.5) can be written as

$$L_1(\tilde{\psi}, \phi) = F_{\omega^\alpha - f}(\phi). \tag{6.8}$$

And similarly for (6.4), we have

$$L_\alpha(\psi^\alpha, \phi) = F_{-\tilde{\psi}}(\phi). \tag{6.9}$$

We discretise Eqs. (6.8) and (6.9) using a continuous Galerkin (CG) discretisation scheme.

Let δ be the discretisation parameter, and let \mathcal{D}_δ denote a space filling triangulation of the domain, that consists of geometry-conforming non-overlapping elements. Define the approximation space

$$W_\delta^k(\mathcal{D}) := \left\{ \phi_\delta \in W^1(\mathcal{D}) : \phi_\delta \in C(\mathcal{D}), \phi_\delta|_K \in \Pi^k(K) \text{ each } K \in \mathcal{D}_\delta \right\}. \tag{6.10}$$

in which $C(\mathcal{D})$ is the space of continuous functions on \mathcal{D} , and $\Pi^k(K)$ denotes the space of polynomials of degree at most k on element $K \in \mathcal{D}_\delta$.

For (6.8), given $f_\delta \in W_\delta^k(\mathcal{D})$ and $\omega_\delta \in V_\delta^k(\mathcal{D})$ (see (6.13) for the definition of $V_\delta^k(\mathcal{D})$), our numerical approximation is the solution $\tilde{\psi}_\delta \in W_\delta^k(\mathcal{D})$ that satisfies

$$L_1(\tilde{\psi}_\delta, \phi_\delta) = F_{\omega_\delta - f_\delta}(\phi_\delta) \tag{6.11}$$

for all test functions $\phi_\delta \in W_\delta^k(\mathcal{D})$. Then, using $\tilde{\psi}_\delta$, our numerical approximation of ψ^α is the solution $\psi_\delta^\alpha \in W_\delta^k(\mathcal{D})$ that satisfies

$$L_\alpha(\psi_\delta^\alpha, \phi_\delta) = F_{-\tilde{\psi}_\delta}(\phi_\delta) \tag{6.12}$$

for all test functions $\phi_\delta \in W_\delta^k(\mathcal{D})$.

For a detailed exposition of the numerical algorithms that solves the discretised problems (6.11) and (6.12), we point the reader to Gibson et al. (2019); Brenner et al. (2008).

6.1.2 Hyperbolic Equations

We choose to discretise the hyperbolic buoyancy (2.12) and potential vorticity (2.13) equations using a discontinuous Galerkin (DG) scheme. For a detailed exposition of DG methods, we refer the interested reader to Hesthaven and Warburton (2007).

Define the DG approximation space, denoted by $V_\delta^k(\mathcal{D})$, to be the element-wise polynomial space,

$$V_\delta^k(\mathcal{D}) = \left\{ v_\delta \in L^2(\mathcal{D}) \mid \forall K \in \mathcal{D}_\delta, \exists \phi_\delta \in \Pi^k(K) : v_\delta|_K = \phi_\delta|_K \right\}. \tag{6.13}$$

We look to approximate b^α and ω^α in the space $V_\delta^k(\mathcal{D})$. Essentially, this means our approximations of b^α and ω^α are each an direct sum over the elements in \mathcal{D}_δ . Additional constraints on the numerical fluxes across shared element boundaries are needed to ensure conservation properties and stability. Further, note that $W_\delta^k(\mathcal{D}) \subset V_\delta^k(\mathcal{D})$. This inclusion is also needed for ensuring numerical conservation laws, see Sect. 6.1.3.

For the buoyancy Eq. (2.12), we obtain the following variational formulation

$$\langle \partial_t b^\alpha, v_\delta \rangle_K = \langle b^\alpha \mathbf{u}^\alpha, \nabla v_\delta \rangle_K - \langle b^\alpha \mathbf{u}^\alpha \cdot \hat{\mathbf{n}}, v_\delta \rangle_{\partial K}, \quad K \in \mathcal{D}_\delta \tag{6.14}$$

where $v_\delta \in V_\delta^k(\mathcal{D})$ is any test function, ∂K denotes the boundary of K , and $\hat{\mathbf{n}}$ denotes the unit normal vector to ∂K . Let b_δ^α be the approximation of b^α in V_δ^k , and let

$\mathbf{u}_\delta^\alpha = \nabla^\perp \psi_\delta^\alpha$ for $\psi_\delta^\alpha \in W_\delta^k$. Our discretised buoyancy equation over each element is given by

$$\langle \partial_t b_\delta^\alpha, v_\delta \rangle_K = \langle b_\delta^\alpha \mathbf{u}_\delta^\alpha, \nabla v_\delta \rangle_K - \langle b_\delta^\alpha \mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}}, v_\delta \rangle_{\partial K}, \quad K \in \mathcal{D}_\delta. \tag{6.15}$$

Similarly, let $\omega_\delta^\alpha \in V_\delta^k(\mathcal{D})$ be the approximation of ω^α , and let $h_\delta \in W_\delta^k$. We obtain the following discretised variational formulation that corresponds to (2.13),

$$\begin{aligned} \langle \partial_t \omega_\delta^\alpha, v_\delta \rangle_K &= \langle (\omega_\delta^\alpha - b_\delta^\alpha) \mathbf{u}_\delta^\alpha, \nabla v_\delta \rangle_K - \langle (\omega_\delta^\alpha - b_\delta^\alpha) \mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}}, v_\delta \rangle_{\partial K} \\ &\quad - \frac{1}{2} \langle \nabla \cdot (b_\delta^\alpha \nabla^\perp h_\delta), v_\delta \rangle_K, \quad K \in \mathcal{D}_\delta, \end{aligned} \tag{6.16}$$

for test function $v_\delta \in V_\delta^k(\mathcal{D})$.

At this point, we only have the discretised problem on single elements. To obtain the global approximation, we sum over all the elements in \mathcal{D}_δ . In doing so, the ∂K terms in (6.15) and (6.16) must be treated carefully. Let ∂K_{ext} denote the part of cell boundary that is contained in $\partial \mathcal{D}$. Let ∂K_{int} denote the part of the cell boundary that is contained in the interior of the domain $\mathcal{D} \setminus \partial \mathcal{D}$. On ∂K_{ext} we simply impose the PDE boundary conditions. However, on ∂K_{int} we need to consider the contribution from each of the neighbouring elements. By choice, the approximant $\psi_\delta^\alpha \in W_\delta^k$ is continuous on ∂K . And since

$$\mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}} = \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}} = -\nabla \psi_\delta^\alpha \cdot \hat{\boldsymbol{\tau}} = -\frac{d\psi_\delta^\alpha}{d\hat{\boldsymbol{\tau}}}, \tag{6.17}$$

where $\hat{\boldsymbol{\tau}}$ denotes the unit tangential vector to ∂K , $\mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}}$ is also continuous. This means $\mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}}$ in (6.15) (also in (6.16)) is single valued. However, due to the lack of global continuity constraint in the definition of $V_k^\alpha(\mathcal{D})$, b_δ^α and ω_δ^α are multi-valued on ∂K_{int} . Thus, as our approximation of ω^α and b^α over the whole domain is the sum over $K \in \mathcal{D}_\delta$, we have to constrain the flux on the set $(\bigcup_{K \in \mathcal{D}_\delta} \partial K) \setminus \partial \mathcal{D}$. This is done using appropriately chosen numerical flux fields in the boundary terms of (6.15) and (6.16).

Let $v^- := \lim_{\epsilon \uparrow 0} v(\mathbf{x} + \epsilon \hat{\mathbf{n}})$ and $v^+ := \lim_{\epsilon \downarrow 0} v(\mathbf{x} + \epsilon \hat{\mathbf{n}})$, for $\mathbf{x} \in \partial K$, be the inside and outside (with respect to a fixed element K) values, respectively, of a function v on the boundary. Let \hat{f} be a numerical flux function that satisfies the following properties:

(i) consistency

$$\hat{f}(v, v, \mathbf{u}_\delta \cdot \hat{\mathbf{n}}) = v \mathbf{u}_\delta \cdot \hat{\mathbf{n}} \tag{6.18}$$

(ii) conservative

$$\hat{f}(v^+, v^-, \mathbf{u}_\delta \cdot \hat{\mathbf{n}}) = -\hat{f}(v^-, v^+, -\mathbf{u}_\delta \cdot \hat{\mathbf{n}}) \tag{6.19}$$

(iii) L^2 stable in the entropy norm with respect to the buoyancy equation, see (Bersens et al., 2006, Section 6).

With such an \hat{f} , we replace $b_\delta^\alpha \mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}}$ by the numerical flux $\hat{f}(b_\delta^{\alpha,+}, b_\delta^{\alpha,-}, \mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}})$ in (6.15). Similarly, in (6.16), we replace $(\omega_\delta^\alpha - b_\delta^\alpha) \mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}}$ by $\hat{f}((\omega_\delta^\alpha - b_\delta^\alpha)^+, (\omega_\delta^\alpha - b_\delta^\alpha)^-, \mathbf{u}_\delta^\alpha \cdot \hat{\mathbf{n}})$.

Remark 6.1 For a general nonlinear conservation law, one has to solve what is called the *Riemann problem* for the numerical flux, see (Hesthaven and Warburton 2007) for details. In our setup, we use the following local Lax–Friedrichs flux, which is an approximate Riemann solver,

$$\hat{f}(v^+, v^-, \mathbf{u}_\delta \cdot \hat{\mathbf{n}}) = \mathbf{u}_\delta \cdot \hat{\mathbf{n}}\{\{v\}\} - \frac{|\mathbf{u}_\delta \cdot \hat{\mathbf{n}}|}{2} \llbracket v \rrbracket \tag{6.20}$$

where

$$\{\{v\}\} := \frac{1}{2}(v^- + v^+), \quad \llbracket v \rrbracket := \hat{\mathbf{n}}^- v^- + \hat{\mathbf{n}}^+ v^+. \tag{6.21}$$

Finally, our goal is to find $b_\delta^\alpha, \omega_\delta^\alpha \in V_\delta^k(\mathcal{D})$ such that for all $v_\delta \in V_\delta^k(\mathcal{D})$ we have

$$\begin{aligned} & \sum_{K \in \mathcal{D}_\delta} \langle \partial_t b_\delta^\alpha, v_\delta \rangle_K \\ &= \sum_{K \in \mathcal{D}_\delta} \{ \langle b_\delta^\alpha \nabla^\perp \psi_\delta^\alpha, \nabla v_\delta \rangle_K - \langle \hat{f}_{b^\alpha}(b_\delta^{\alpha,+}, b_\delta^{\alpha,-}, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}, v_\delta^-) \rangle_{\partial K} \}, \end{aligned} \tag{6.22}$$

$$\begin{aligned} \sum_{K \in \mathcal{D}_\delta} \langle \partial_t \omega_\delta^\alpha, v_\delta \rangle_K &= \sum_{K \in \mathcal{D}_\delta} \{ \langle (\omega_\delta^\alpha - b_\delta^\alpha) \nabla^\perp \psi_\delta^\alpha, \nabla v_\delta \rangle_K - \frac{1}{2} \langle \nabla \cdot (b_\delta^\alpha \nabla^\perp h_\delta), v_\delta \rangle_K \\ &\quad - \langle \hat{f}_{\omega^\alpha}((\omega_\delta^\alpha - b_\delta^\alpha)^+, (\omega_\delta^\alpha - b_\delta^\alpha)^-, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}, v_\delta^-) \rangle_{\partial K} \} \end{aligned} \tag{6.23}$$

with $\psi_\delta^\alpha \in W_\delta^k(\mathcal{D})$ being the numerical approximation to the stream function.

Remark 6.2 In (6.22) and (6.23) we do not explicitly distinguish ∂K_{ext} and ∂K_{int} because for the boundary conditions (6.1), the ∂K_{ext} terms vanish.

6.1.3 Numerical Conservation

Conservation properties of the TQG system was first shown in Holm et al. (2021). More specifically, the TQG system conserves energy, and an infinite family of quantities called Casimirs. Proposition 6.3 below describes the conservation properties of the α -TQG system. We note that the form of the conserved energy and Casimirs are the same as that of the TQG system. Although the result is stated for the system with boundary conditions, it is easy to show that the same result holds for \mathbb{T}^2 .

Proposition 6.3 (α -TQG conserved quantities) *On a bounded domain \mathcal{D} with boundary $\partial\mathcal{D}$, consider the α -TQG system (2.12)–(2.14) with boundary conditions*

$$\begin{aligned} \hat{\mathbf{n}} \cdot \mathbf{u}^\alpha &= 0, \quad \hat{\mathbf{n}} \times \nabla b^\alpha = 0, \quad \frac{d}{dt} \int_{\partial\mathcal{D}} \nabla \psi^\alpha \cdot \hat{\mathbf{n}} = 0, \\ \frac{d}{dt} \int_{\partial\mathcal{D}} \nabla \Delta \psi^\alpha \cdot \hat{\mathbf{n}} &= 0, \quad \hat{\mathbf{n}} \times \nabla \Delta \psi^\alpha = 0 \end{aligned} \tag{6.24}$$

We have

$$E^\alpha(t) := -\frac{1}{2} \int_{\mathcal{D}} \{ (\omega^\alpha - f) \psi^\alpha + h b^\alpha \} \, dx, \tag{6.25}$$

and

$$C_{\Psi, \Phi}^\alpha(t) := \int_{\mathcal{D}} \{\Psi(b^\alpha) + \omega^\alpha \Phi(b^\alpha)\} \, d\mathbf{x}, \quad \forall \Psi, \Phi \in C^\infty \tag{6.26}$$

are conserved, i.e. $d_t E^\alpha(t) = 0$ and $d_t C_{\Psi, \Phi}^\alpha(t) = 0$.

Proof For the energy $E^\alpha(t)$, we obtain

$$\begin{aligned} \frac{d}{dt} E^\alpha(t) &= -\frac{1}{2} \frac{d}{dt} \int_{\mathcal{D}} \{\psi^\alpha(\Delta - 1)(1 - \alpha\Delta)\psi^\alpha + hb^\alpha\} \, d\mathbf{x} \\ &= \int_{\mathcal{D}} \{-\psi^\alpha \partial_t \omega^\alpha - \frac{1}{2} h \partial_t b^\alpha\} \, d\mathbf{x}. \end{aligned}$$

Substituting $\partial_t \omega^\alpha$ and $\partial_t b^\alpha$ using (2.12) and (2.13), the result then follows from direct calculations.

Similarly for the Casimirs, we obtain the result by directly evaluating $d_t C_{\Psi, \Phi}^\alpha(t)$. □

Remark 6.4 In (6.24), except for the integral Neumann boundary conditions, we effectively have Dirichlet boundary conditions for ψ^α , b^α and $\Delta\psi^\alpha$. The integral Neumann boundary conditions can be viewed as imposing the Kelvin theorem on the boundary—consider

$$\frac{d}{dt} \int_{\partial\mathcal{D}} \nabla(\psi^\alpha - \alpha\Delta\psi^\alpha) \cdot \hat{\mathbf{n}} = 0, \tag{6.27}$$

and apply the divergence theorem.

In our FEM discretisation, we impose only Dirichlet boundary conditions. If the integral Neumann boundary conditions are not imposed, then for Proposition 6.3 to hold, we necessarily require ψ^α and $\Delta\psi^\alpha$ on $\partial\mathcal{D}$ to be zero.

We now analyse the conservation properties of our spatially discretised α -TQG system. Let $\langle \cdot, \cdot \rangle_{H^1(\mathcal{D})}$ be the $H^1(\mathcal{D})$ inner product. Define the numerical total energy

$$E_\delta^\alpha(t) = \frac{1}{2} \langle \psi_\delta^\alpha, \tilde{\psi}_\delta \rangle_{H^1(\mathcal{D})} - \frac{1}{2} \langle h_\delta, b_\delta^\alpha \rangle_{\mathcal{D}}, \tag{6.28}$$

in which $\psi_\delta^\alpha, \tilde{\psi}_\delta \in W_\delta^\alpha(\mathcal{D})$ and $b_\delta^\alpha \in V_\delta^k(\mathcal{D})$ are the numerical approximations of ψ^α , $\tilde{\psi}$ and b^α , respectively. Define the numerical Casimir functional

$$C_\delta^\alpha(t; \Psi, \Phi) := \langle \Psi(b_\delta^\alpha), 1 \rangle_{\mathcal{D}} + \langle \omega_\delta^\alpha \Phi(b_\delta^\alpha) \rangle_{\mathcal{D}}, \quad \Psi, \Phi \in C^\infty. \tag{6.29}$$

Lemma 6.5 *With boundary conditions (6.1)*

$$\frac{d}{dt} E_\delta^\alpha(t) = 0 \tag{6.30}$$

for $\alpha = 0$.

In other words, the semi-discrete discretisation conserves energy in the TQG case.

Proof We consider d/d_t evaluated at an arbitrary fixed value t_0 . First note that when $\alpha = 0$, we have $\psi_\delta^\alpha = \psi_\delta$. Then, from (6.6) we have

$$2E_\delta^\alpha(t) = L_1(\psi_\delta^\alpha, \psi_\delta^\alpha) - \langle h_\delta, b_\delta^\alpha \rangle_\Omega \tag{6.31}$$

For the first term in (6.31), following (6.11) we have

$$L_1(\partial_t \psi_\delta, \phi_\delta)|_{t=t_0} = - \sum_{K \in \mathcal{D}_\delta} \langle \partial_t \omega_\delta^\alpha, \phi_\delta \rangle_K |_{t=t_0}, \quad \forall \phi_\delta \in W_\delta^k(\mathcal{D}). \tag{6.32}$$

Thus, substituting in the discretised equation (6.23) for ω^α , we obtain

$$\begin{aligned} -L_1(\partial_t \psi_\delta, \phi_\delta)|_{t=t_0} &= \sum_{K \in \mathcal{D}_\delta} \left[\langle (\omega_\delta^\alpha - b_\delta^\alpha) \nabla^\perp \psi_\delta^\alpha, \nabla \phi_\delta \rangle_K - \frac{1}{2} \langle \nabla \cdot (b_\delta^\alpha \nabla^\perp h_\delta), \phi_\delta \rangle_K \right. \\ &\quad \left. - \langle \hat{f}_{\omega^\alpha}((\omega_\delta^\alpha - b_\delta^\alpha)^+, (\omega_\delta^\alpha - b_\delta^\alpha)^-, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}), \phi_\delta^- \rangle_{\partial K} \right] |_{t=t_0}, \end{aligned} \tag{6.33}$$

in which we can choose $\phi_\delta = [\psi_\delta^\alpha]_{t=t_0}$. This choice is consistent with the discretisation space $V_\delta^k(\mathcal{D})$ for ω^α , since $W_\delta^k(\mathcal{D}) \subset V_\delta^k(\mathcal{D})$.

Similarly, for the second term in (6.31), using (6.22) we obtain

$$\begin{aligned} &\frac{d}{dt} \Big|_{t=t_0} \langle b_\delta^\alpha, v_\delta \rangle_{\mathcal{D}} \\ &= \sum_{K \in \mathcal{D}_\delta} \left[\langle b_\delta^\alpha \nabla^\perp \psi_\delta^\alpha, \nabla v_\delta \rangle_K - \langle \hat{f}_{b^\alpha}(b_\delta^{\alpha,+}, b_\delta^{\alpha,-}, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}), v_\delta^- \rangle_{\partial K} \right] |_{t=t_0} \end{aligned} \tag{6.34}$$

where we can choose $v_\delta = h_\delta$.

Putting together (6.31), (6.33) and (6.34), and substituting in h_δ and $[\psi_\delta^\alpha]_{t=t_0}$ for the arbitrary choices, we obtain

$$\begin{aligned} 2 \frac{d}{dt} \Big|_{t=t_0} E_\delta^\alpha(t) &= - \sum_{K \in \mathcal{D}_\delta} \left[2 \left\{ \underbrace{\langle (\omega_\delta^\alpha - b_\delta^\alpha) \nabla^\perp \psi_\delta^\alpha, \nabla \psi_\delta^\alpha \rangle_K}_{=0} - \frac{1}{2} \langle \nabla \cdot (b_\delta^\alpha \nabla^\perp h_\delta), \psi_\delta^\alpha \rangle_K \right. \right. \\ &\quad \left. \left. - \langle \hat{f}_{\omega^\alpha}((\omega_\delta^\alpha - b_\delta^\alpha)^+, (\omega_\delta^\alpha - b_\delta^\alpha)^-, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}), \psi_\delta^{\alpha-} \rangle_{\partial K} \right\} \right. \\ &\quad \left. + \left\{ \underbrace{\langle b_\delta^\alpha \nabla^\perp \psi_\delta^\alpha, \nabla h_\delta \rangle_K}_{=-\langle b_\delta^\alpha \nabla^\perp h_\delta, \nabla \psi_\delta^\alpha \rangle_K} - \langle \hat{f}_{b^\alpha}(b_\delta^{\alpha,+}, b_\delta^{\alpha,-}, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}), h_\delta^- \rangle_{\partial K} \right\} \right] |_{t=t_0} \end{aligned} \tag{6.35}$$

$$\begin{aligned} \frac{d}{dt} E_\delta^\alpha(t) &= \sum_{K \in \mathcal{D}_\delta} \left[- \langle \hat{f}_{\omega^\alpha}((\omega_\delta^\alpha - b_\delta^\alpha)^+, (\omega_\delta^\alpha - b_\delta^\alpha)^-, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}), \psi_\delta^{\alpha-} \rangle_{\partial K} \right. \\ &\quad \left. - \frac{1}{2} \langle \hat{f}_{b^\alpha}(b_\delta^{\alpha,+}, b_\delta^{\alpha,-}, \nabla^\perp \psi_\delta^\alpha \cdot \hat{\mathbf{n}}), h_\delta^- \rangle_{\partial K} \right] |_{t=t_0} \end{aligned} \tag{6.36}$$

Since ψ_δ and h_δ are continuous across element boundaries, we have $\psi_\delta^+ = \psi_\delta^-$ and $h_\delta^+ = h_\delta^-$. Thus, if the numerical fluxes \hat{f}_{ω^α} and \hat{f}_{b^α} satisfy the conservative property (6.19), the sum of flux terms is zero. \square

Remark 6.6 In order for our semi-discrete discretisation to conserve numerical energy when $\alpha > 0$, more regularity is required for the approximation space for ψ_δ^α . Additionally, the current scheme coupled with the time stepping algorithm does not conserve Casimirs. For future work, we look to explore Casimir-conserving schemes for the model.

6.1.4 Time Stepping

To discretise the time derivative, we use the strong stability preserving Runge–Kutta of order 3 (SSPRK3) scheme, see (Hesthaven and Warburton 2007). Writing the finite element spatial discretisation formally as

$$\partial_t b_\delta^\alpha = f_\delta(b_\delta^\alpha) \tag{6.37}$$

where f_δ is the discretisation operator that follows from (6.22), and

$$\partial_t \omega = g_\delta(\omega, b) \tag{6.38}$$

where g_δ is the discretisation operator that follows from (6.23). Let $b_\delta^{\alpha,n}$ and $\omega_\delta^{\alpha,n}$ denote the approximation of b_δ^α and ω_δ^α at time step t_n . The SSPRK3 time discretisation is as follows

$$b^{(1)} = b_\delta^{\alpha,n} + \Delta t f_\delta(b_\delta^{\alpha,n}) \tag{6.39a}$$

$$\omega^{(1)} = \omega_\delta^{\alpha,n} + \Delta t g_\delta(\omega_\delta^{\alpha,n}, b_\delta^{\alpha,n}) \tag{6.39b}$$

$$b^{(2)} = \frac{3}{4} b_\delta^{\alpha,n} + \frac{1}{4} (b^{(1)} + \Delta t f_\delta(b^{(1)})) \tag{6.39c}$$

$$\omega^{(2)} = \frac{3}{4} \omega_\delta^{\alpha,n} + \frac{1}{4} (\omega^{(1)} + \Delta t g_\delta(\omega^{(1)}, b^{(1)})) \tag{6.39d}$$

$$b_\delta^{\alpha,n+1} = \frac{1}{3} b_\delta^{\alpha,n} + \frac{2}{3} (b^{(2)} + \Delta t f_\delta(b^{(2)})) \tag{6.39e}$$

$$\omega_\delta^{\alpha,n+1} = \frac{1}{3} \omega_\delta^{\alpha,n} + \frac{2}{3} (\omega^{(2)} + \Delta t g_\delta(\omega^{(2)}, b^{(2)})) \tag{6.39f}$$

where $\Delta t = t_{n+1} - t_n$ each n .

6.2 α -TQG Linear Thermal Rossby Wave Stability Analysis

A dispersion relation for the TQG system was derived in Holm et al. (2021). There, the authors showed that the linear thermal Rossby waves of the TQG system possess high wavenumber instabilities. More specifically, the Doppler-shifted phase speed of these waves becomes complex at sufficiently high wavenumbers. However, the growth rate of

the instability decreases to zero as $|\mathbf{k}|^{-1}$, in the limit that $|\mathbf{k}| \rightarrow \infty$. Consequently, the TQG dynamics is linearly well-posed. That is, the TQG solution depends continuously on initial conditions. In this subsection, using the same equilibrium state as in Holm et al. (2021), we derive a dispersion relation for the thermal Rossby wave solutions of the linearised α -TQG system. Again the Doppler-shifted phase speed of these waves becomes complex at sufficiently high wavenumbers. However, in this case, the growth rate of the instability decreases to zero as $|\mathbf{k}|^{-2}$, in the limit that $|\mathbf{k}| \rightarrow \infty$. In the limit that α tends to zero, one recovers the TQG dispersion relation derived in Holm et al. (2021).

For the reader’s convenience, we repeat the equations solved by the pair of variables $(b^\alpha, \omega^\alpha)$ in the α -TQG system in (2.12)–(2.13), as formulated now in vorticity-streamfunction form with fluid velocity given by $\mathbf{u}^\alpha = \nabla^\perp \psi^\alpha$. This formulation is given by

$$\frac{\partial}{\partial t} b^\alpha + J(\psi^\alpha, b^\alpha) = 0, \tag{6.40}$$

$$\frac{\partial}{\partial t} \omega^\alpha + J(\psi^\alpha, \omega^\alpha - b^\alpha) = -\frac{1}{2} J(h, b^\alpha), \tag{6.41}$$

$$\omega^\alpha = (\Delta - 1)(1 - \alpha\Delta)\psi^\alpha + f. \tag{6.42}$$

Here, as before, $J(a, b) = \nabla^\perp a \cdot \nabla b = a_x b_y - b_x a_y$ denotes the Jacobian of any smooth functions a and b defined on the (x, y) plane.

Equilibrium states of the α -TQG system in the equations (6.40)–(6.42) satisfy $J(\psi_e^\alpha, \omega_e^\alpha) = J(\psi_e^\alpha, b_e^\alpha) = J(h, b_e^\alpha) = 0$. The equilibrium TQG state considered in Holm et al. (2021) was given via the specification of the following gradient fields

$$\begin{aligned} \nabla \psi_e^\alpha &= -U \hat{\mathbf{y}}, \quad \nabla \omega_e^\alpha = (U - \beta) \hat{\mathbf{y}}, \quad \nabla f = -\beta \hat{\mathbf{y}} \\ \nabla b_e^\alpha &= -B \hat{\mathbf{y}}, \quad \nabla h = -H \hat{\mathbf{y}} \end{aligned}$$

where we have taken $f = 1 - \beta y$ and the equilibrium parameters $U, \beta, B, H \in \mathbb{R}$ are all constants. Linearising the α -TQG system around the steady state produces the following evolution equations for the perturbations ω', b' and ψ'

$$\omega'_t + U \omega'_x + (U + B - \beta) \psi'_x = (U - H/2) b'_x \tag{6.43a}$$

$$b'_t + U b'_x - B \psi'_x = 0 \tag{6.43b}$$

$$\left((1 - \alpha\Delta)(1 - \Delta) \right) \psi' = -\omega'. \tag{6.43c}$$

Since these equations are linear with constant coefficients, one has the plane wave solutions

$$\omega' = e^{i(\mathbf{k}\cdot\mathbf{x} - \nu t)} \hat{\omega}, \quad b' = e^{i(\mathbf{k}\cdot\mathbf{x} - \nu t)} \hat{b}, \quad \psi' = e^{i(\mathbf{k}\cdot\mathbf{x} - \nu t)} \hat{\psi}.$$

Further, from (6.43c) we obtain

$$-\hat{\omega} = (|\mathbf{k}|^2 + 1)(\alpha|\mathbf{k}|^2 + 1)\hat{\psi}.$$

Substituting these solutions into the linearised equations, we have

$$\begin{aligned} (v - kU)\hat{\omega} - k(U + B - \beta)\hat{\psi} &= -k(U - H/2)\hat{b} \\ (v - kU)\hat{b} &= -kB\hat{\psi} \end{aligned}$$

where k is the first component of \mathbf{k} . From the above we obtain a quadratic formula for the Doppler-shifted phase speed $C = C(\alpha) := (v(\alpha) - kU) / k$

$$C^2(|\mathbf{k}|^2 + 1)(\alpha|\mathbf{k}|^2 + 1) + CX + Y = 0.$$

where $X := U + B - \beta$ and $Y := (U - H/2)B$. Thus, the dispersion relation for the thermal Rossby wave for TQG possesses two branches, corresponding to two different phase velocities,

$$C(\alpha) = \frac{-X \pm \sqrt{X^2 - 4Y(|\mathbf{k}|^2 + 1)(\alpha|\mathbf{k}|^2 + 1)}}{2(|\mathbf{k}|^2 + 1)(\alpha|\mathbf{k}|^2 + 1)}. \tag{6.44}$$

Upon setting $\alpha = 0$ in (6.44), we recover the Doppler-shifted phase speed of thermal Rossby waves for the TQG system in Holm et al. (2021). When $\alpha = 0$, $B = 0$, and $U = 0$ in (6.44), the remaining dispersion relation differs slightly from the dispersion relation for QG Rossby waves in non-dimensional variables. This is because the Casimirs in the Hamiltonian formulation of the TQG system differ from the those of standard QG.

As discussed in Holm et al. (2021) for the TQG system, when $Y > 0$, the Doppler shifted phase velocity in (6.44) for the linearised wave motion becomes complex at high wavenumber $|\mathbf{k}| \gg 1$; namely for $(|\mathbf{k}|^2 + 1)(\alpha|\mathbf{k}|^2 + 1) \geq X^2/(4Y)$, for both branches of the dispersion relation. However, the growth rate of the instability found from the imaginary part of $C(\alpha)$ in (6.44) decays to zero as $O(|\mathbf{k}|^{-2})$, for $|\mathbf{k}| \gg 1$. Indeed, the linear stability analysis leading to $C(\alpha)$ predicts that the maximum growth rate occurs at a finite wavenumber $|\mathbf{k}|_{\max}$ which depends on the value of α , beyond which the growth rate of the linearised TQG wave amplitude falls rapidly to zero. One would expect that simulated solutions of α -TQG would be most active at the length-scale corresponding to $|\mathbf{k}|_{\max}$, at which the linearised thermal Rossby waves are the most unstable. If this maximum activity length-scale is near the grid truncation size, then numerical truncation errors could cause additional numerical stability issues! We have experienced such problems during our testing of the numerical algorithm for the TQG system. Even for equilibrium solution ‘‘sanity’’ check tests, we have found that unless the time step was taken to be incredibly small, high wavenumber truncation errors have caused our numerical solutions to eventually blow up.

The magnitude of α controls the unstable growth rate of linearised TQG waves at asymptotically high wavenumbers. That is, the presence of α regularises the wave

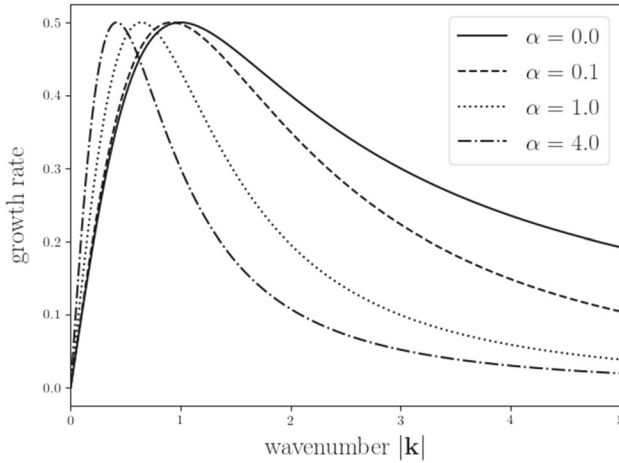


Fig. 4 An example of the growth-rate of linearised TQG waves as determined from the imaginary part of $C(\alpha)$ in Eq. (6.44), plotted for different values of α . We observe that increasing α shifts the wavenumber at the maximum growth rate to lower values

activity at high wavenumbers, see Fig. 4. Thus, from the perspective of numerics, the α regularisation is available to control numerical problems that may arise from the inherent model instability at high wavenumbers, without the need for additional dissipation terms in the equation.

6.3 Numerical Example

The numerical set-up we consider for this paper is as follows. The spatial domain \mathcal{D} is an unit square with doubly periodic boundaries. We chose to discretise \mathcal{D} using a grid that consists of 256×256 cells, i.e. $\text{cardinality}(\mathcal{D}_\delta) = 256 \times 256$. This was the maximum resolution we could computationally afford, to obtain results over a reasonable amount of time.

We computed α -TQG solutions for the following values of α — $\frac{1}{16^2}, \frac{1}{32^2}, \frac{1}{64^2}, \frac{1}{128^2}, \frac{1}{180^2}, \frac{1}{220^2}, \frac{1}{256^2}$ and 0. Note when $\alpha = 0$ we get the TQG system. For all cases, the following initial conditions were used,

$$\omega(0, x, y) = \sin(8\pi x) \sin(8\pi y) + 0.4 \cos(6\pi x) \cos(6\pi y) + 0.3 \cos(10\pi x) \cos(4\pi y) + 0.02 \sin(2\pi y) + 0.02 \sin(2\pi x) \tag{6.45}$$

$$b(0, x, y) = \sin(2\pi y) - 1, \tag{6.46}$$

see Fig. 5 for illustrations, as well as the following bathymetry and rotation fields

$$h(x, y) = \cos(2\pi x) + 0.5 \cos(4\pi x) + \frac{1}{2} \cos(6\pi x) \tag{6.47}$$

$$f(x, y) = 0.4 \cos(4\pi x) \cos(4\pi y). \tag{6.48}$$

For time stepping, we used $\Delta t = 0.0005$ in all cases to facilitate comparisons. This choice satisfies the CFL condition for the $\alpha = 0$ case.

We computed each solution for 5000 time steps.¹ Figures 6, 7 and 8 show α -TQG solution snapshots of buoyancy, potential vorticity and velocity magnitude, respectively, at the 1280'th and 2600'th time steps, for α values 0 , $\frac{1}{128^2}$, $\frac{1}{64^2}$ and $\frac{1}{16^2}$. The interpretation of the regularisation parameter α is that its square root value corresponds to the fraction of the domain's length scale that get regularised. At the 1280'th time step, the flows are in early spin-up phase. At the 2600'th time step, although the flows are still in spin-up phase, much more flow features have developed. The subfigures illustrate via comparisons, how α controls the development of small scale features and instabilities. Increasing α leads to more regularisation at larger scales.

In view of Proposition 4.1, we investigated numerically the convergence of α -TQG to TQG using our numerical setup. Consider the relative error between TQG buoyancy and α -TQG buoyancy,

$$e_b(t, \alpha) := \frac{\|b(t) - b^\alpha(t)\|_{H^1}}{\|b(t)\|_{H^1}} \quad (6.49)$$

and the relative error between TQG potential vorticity and α -TQG potential vorticity,

$$e_\omega(t, \alpha) := \frac{\|\omega(t) - \omega^\alpha(t)\|_{L^2}}{\|\omega(t)\|_{L^2}}. \quad (6.50)$$

The norms were chosen in view of Proposition 4.1. Figure 9a and b shows plots of $e_b(t, \alpha)$ and $e_\omega(t, \alpha)$ as functions of time only, for fixed α values $\frac{1}{16^2}$, $\frac{1}{32^2}$, $\frac{1}{64^2}$, $\frac{1}{128^2}$ and $\frac{1}{220^2}$, from the initial time up to and including the 2800'th time step. We observe that starting from 0, the relative errors e_b and e_ω initially increase over time but plateau at around and beyond the 2000'th time step. Up to the 1400'th time step, the plotted relative errors remain less than 1.0 and are arranged in the ascending order of α , i.e. smaller α 's give smaller relative errors.

We note that, if we compare and contrast the relative error results with the solution snapshots of buoyancy and potential vorticity, particularly at the 2600'th time step (shown in Figs. 6b and 7b), we observe “discrepancies”—for lack of a better term—between the results. For example, in Fig. 6b, if we compare the $\alpha = \frac{1}{128^2}$ snapshot with the $\alpha = \frac{1}{16^2}$ one, the former show small scale features that are much closer to those that exist in the reference TQG $\alpha = 0$ solution. However, according to the relative error measurements, the two regularised solutions are more or less equivalent in their differences to the reference TQG solution at time step 2600.

If we fix the value of the time parameter in $e_b(t, \alpha)$ and $e_\omega(t, \alpha)$, and vary α , we can estimate the convergence rate of α -TQG to TQG for our numerical setup. Proposition 4.1 predicts a convergence rate of 1 in α , using the H^1 norm on the buoyancy differences and the L^2 norm on the potential vorticity differences. Note that, in (4.1), the left-hand side evaluates a supremum over a given compact time interval. We see in Fig. 9a and b that the relative errors are monotonic up to around the

¹ Go to <https://youtu.be/a2a4xzf3Pg> for a video of the full simulation.

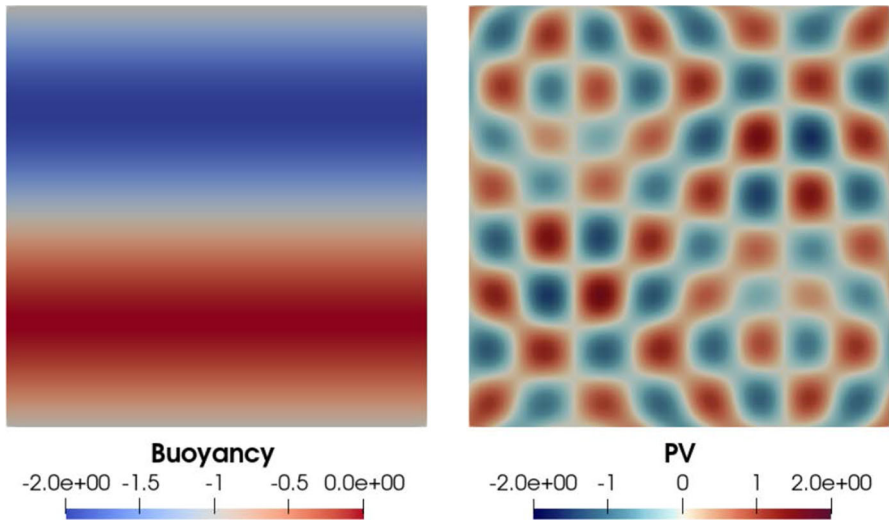


Fig. 5 Initial conditions—buoyancy (6.46) on the left, and potential vorticity (6.45) on the right. The different colours of the PV plot correspond to positive and negative values of PV, which can be interpreted as clockwise and anticlockwise eddies, respectively

1500'th time step mark. So for a given time point T between the initial time and the 1500'th time step, we assume we can evaluate e_b and e_ω at T to estimate the supremum over $[0, T]$.

Figure 10a and b shows plots—in log-log scale—of e_b and e_ω as functions of α only at the 600'th, 800'th, 1000'th and 1200'th time steps, for all the chosen α values. These numbers of time steps correspond to $T = 0.3$, $T = 0.4$, $T = 0.5$ and $T = 0.6$, respectively. We also plotted in Fig. 10a and b linear functions of α to provide reference order 1 slopes. Comparing the results to the reference, we see that the theoretical convergence rate of 1 is attained for $T = 0.3$ (600'th time step), $T = 0.4$ (800'th time step) and $T = 0.5$ (1000'th time step). However for $T = 0.6$ (1200'th time step), the convergence rates are at best order 1/2.

7 Conclusion and Outlook

In conclusion, we have formulated the thermal quasi-geostrophic (TQG) model equations in the regime of approximations relevant to GFD and we have explored their analytical and numerical properties. In respect to the analytical properties, we have shown that the TQG model and its α -regularised version, the α -TQG model, admit unique local strong solutions that are stable in a larger space with weaker norm, that both models have a maximum time of existence, solutions of the latter model converge to solutions of the former model on any time interval in which a solution to the former lives, and we have also determined conditions under which the solution of the α -TQG will blow up.

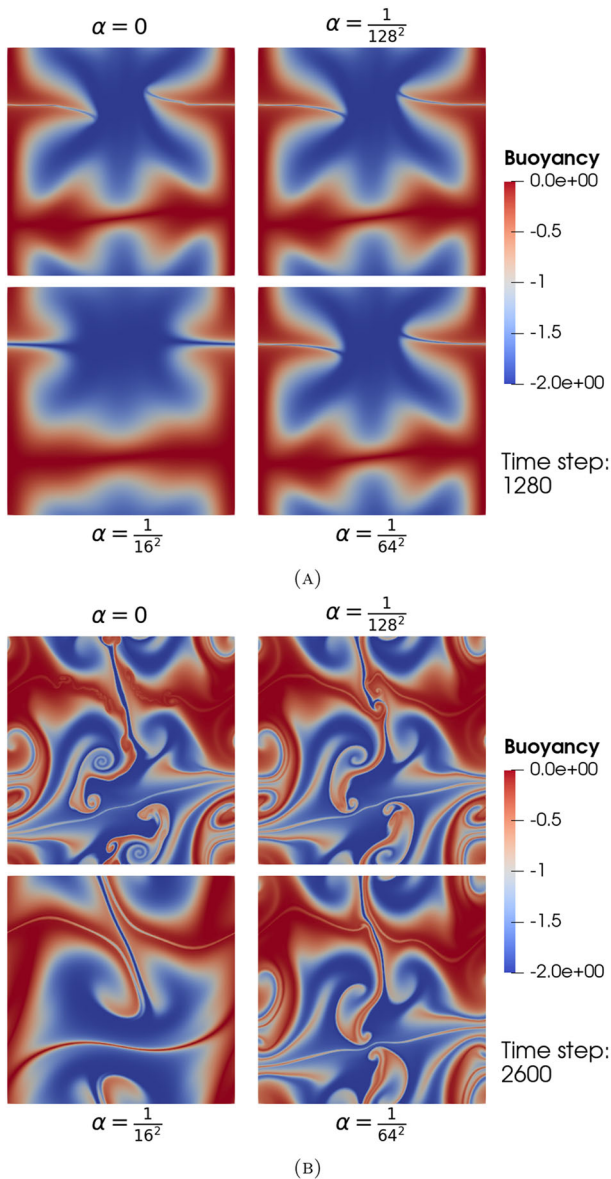


Fig. 6 Comparisons of solution snapshots of the α -TQG buoyancy field that correspond to four different values of α , at two different points in time. Potential vorticity and velocity magnitude snapshots of the same solutions are shown in Figs. 7 and 8, respectively. Shown in each subfigure are the results corresponding to $\alpha = 0$ (top left), $\alpha = \frac{1}{128^2}$ (top right), $\alpha = \frac{1}{64^2}$ (bottom right) and $\alpha = \frac{1}{16^2}$ (bottom left). When $\alpha = 0$, the solution is of the TQG system. Subfigure (A) shows the solutions at the 1280th time step, or equivalently when $t = 0.64$. Subfigure (B) shows the solutions at the 2600th time step, or equivalently when $t = 1.30$. The flows in both subfigures are at different stages of spin-up, with the flow at $t = 1.30$ showing more developed features. The α parameter is interpreted as the fraction of the domain's length-scale squared value at which regularisation is applied. Due to regularisation, the flows do develop differently. Nevertheless, we observe that as α gets smaller, the α -TQG flow features converge to that of the TQG flow

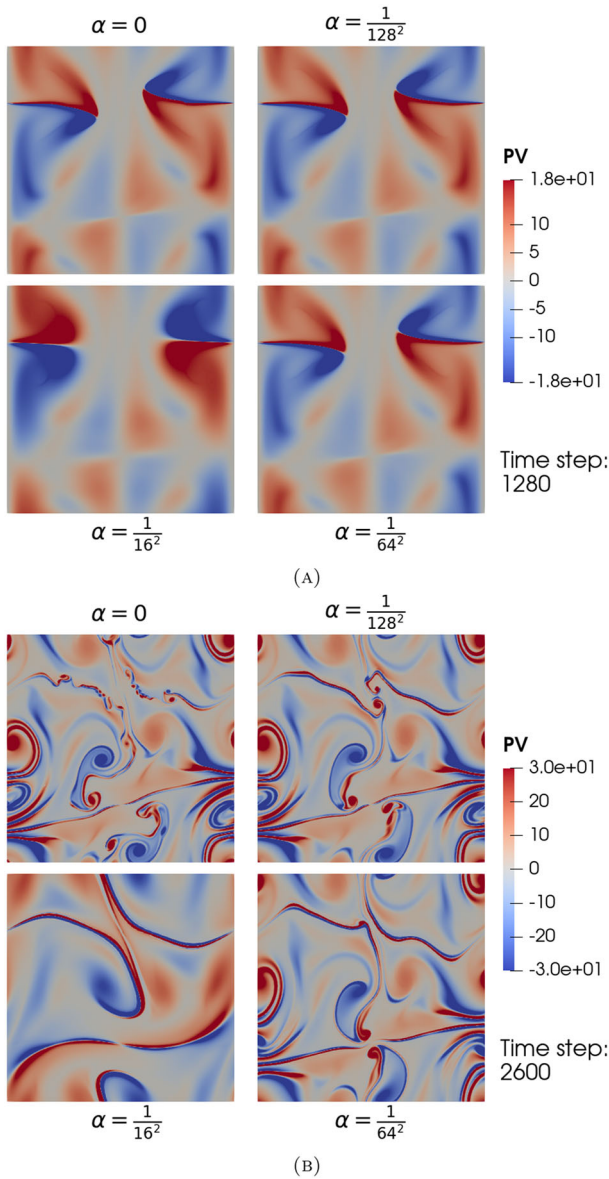


Fig. 7 Comparisons of solution snapshots of the α -TQG potential vorticity field that correspond to four different values of α , at two different points in time. See the caption of Fig. 6 for explanations of the arrangements of these plots. Buoyancy and velocity magnitude snapshots of the same solutions are shown in Figs. 6 and 8, respectively. As in Fig. 5, the different colours of the PV field correspond to positive and negative values of PV, which can be interpreted as clockwise and anticlockwise eddies, respectively. In subfigure (B), we observe more clearly the regularisation effects of α . As α increases, the flows develop in different ways—features at smaller scales no longer develop, and larger features evolve differently as a result. Nevertheless, we observe that as α gets smaller, the α -TQG flow features converge to that of the TQG flow

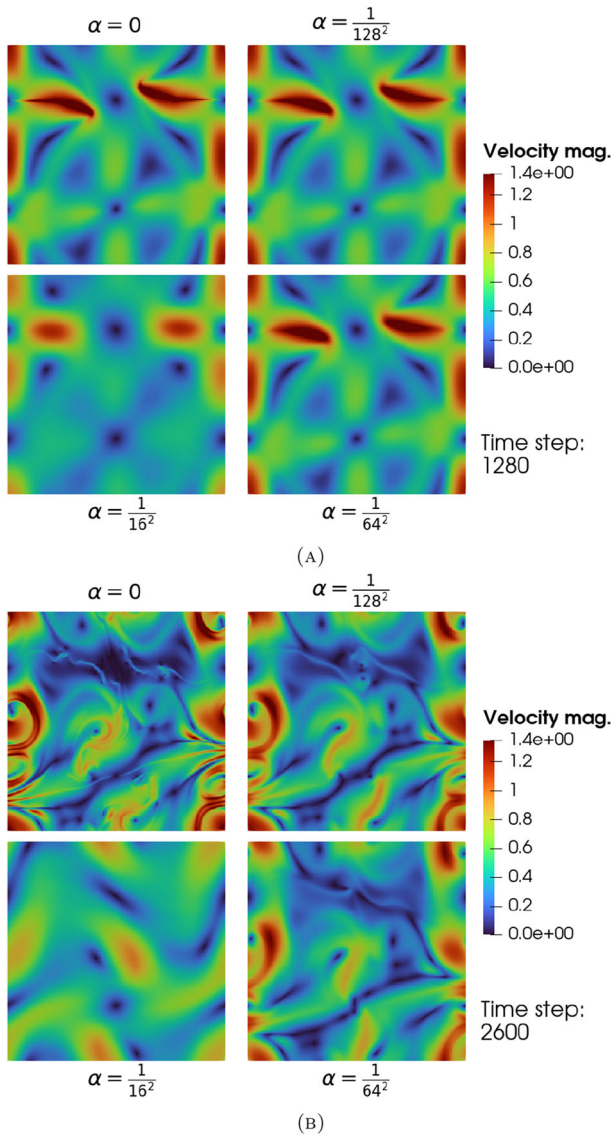


Fig. 8 Comparisons of solution snapshots of the α -TQG velocity magnitudes that correspond to four different values of α , at two different points in time. See the caption of Fig. 6 for explanations of the arrangements of these plots. Buoyancy and potential vorticity snapshots of the same solutions are shown in Figs. 6 and 7, respectively. Using the same scale for colouring, we observe in both subfigures the strength of the colours weaken as α increases. This indicates smoothing of large velocity magnitudes. Additionally, in subfigure (B), we observe how features at smaller scales get smoothed out. In particular, in the $\alpha = \frac{1}{16^2}$ plot, we see that essentially only large scale vortices remain. Hence, considering the velocity field as a part of the system's nonlinear advection operator, these figures help us to better visualise how the flow features developed in Figs. 6 and 7

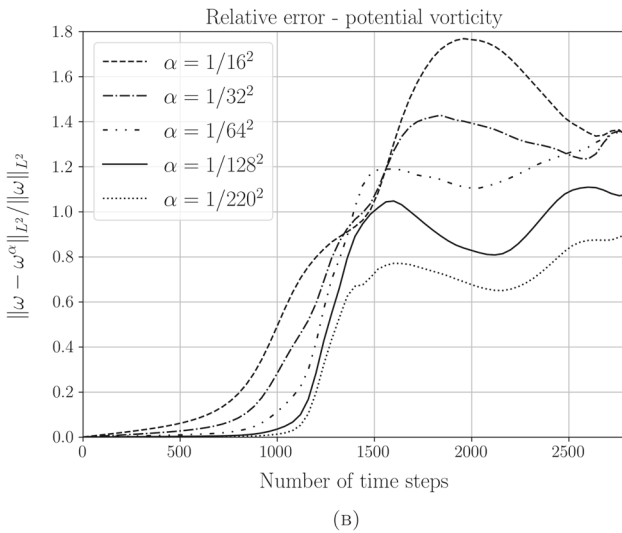
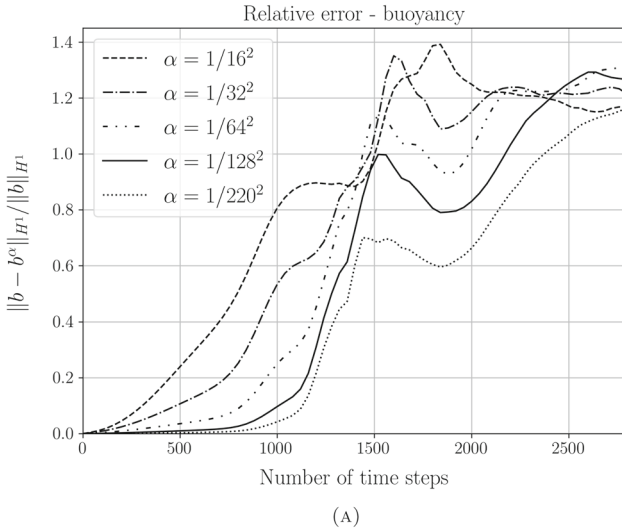


Fig. 9 Subfigures (A) and (B) show, respectively, plots of the relative error functions $e_b(t, \alpha)$ and $e_\omega(t, \alpha)$ as functions of time only, for the fixed α values $\frac{1}{16^2}$, $\frac{1}{32^2}$, $\frac{1}{64^2}$, $\frac{1}{128^2}$ and $\frac{1}{220^2}$. See Eqs. (6.49) and (6.50) for the definitions of e_b and e_ω , respectively. The plots are shown from $t = 0$ up to and including $t = 1.4$, which is equivalent to 2800 time steps using $\Delta t = 0.0005$. We observe that, starting from 0, all plots of e_b and e_ω increase initially, and plateau at around the 2000th time step. Further, up to the 1400th time step the plotted relative errors remain less than 1.0 and are arranged in the ascending order of α . Relating these relative error results at the 2600th time step to the solution snapshots at the same time point (shown in Figs. 6b and 7b), we see that although the snapshots show convergence of flow features, the relative error values suggest all the α -TQG solutions are more or less equally far away from the TQG flow

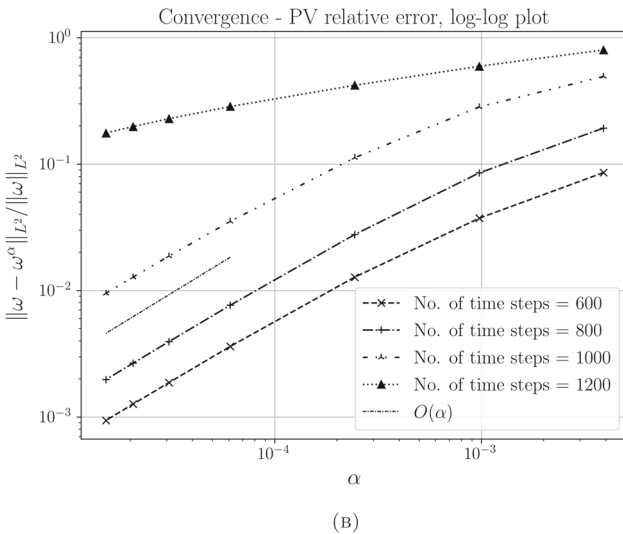
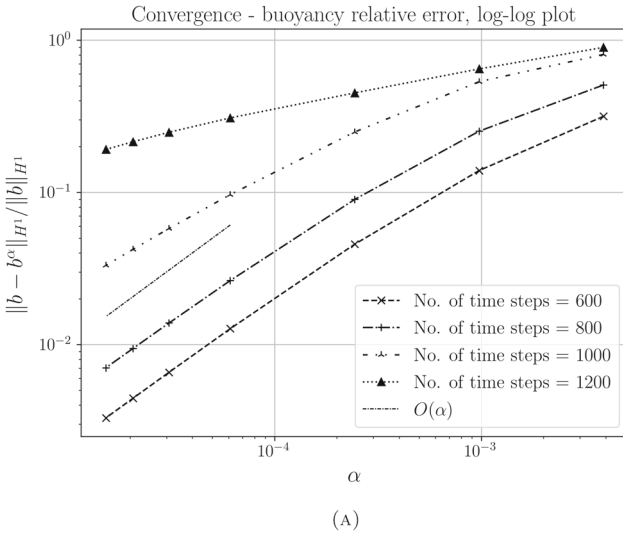


Fig. 10 Subfigures (A) and (B) show, respectively, plots in log-log scale of the relative error functions $e_b(t, \alpha)$, and $e_\omega(t, \alpha)$ as functions of α only, at the fixed time values $t = 0.3$ (equivalently at the 600'th time step), $t = 0.4$ (equivalently at the 800'th time step), $t = 0.5$ (equivalently at the 1000'th time step) and $t = 0.6$ (equivalently at the 1200'th time step). In view of Proposition 4.1, we assumed that, based on the initial monotonicity of the relative errors (see Fig. 9), the supremum over $[0, T]$ in (4.1) can be estimated by evaluating e_b and e_ω at T . Proposition 4.1 predicts that up to a certain time, the rate of convergence of e_b and e_ω to 0 should be no less than order 1 in α . In both subfigures we have plotted, as the reference for comparison, linear functions of α . Thus, comparing slopes to the reference, we see that numerically we get order 1 convergence for $t = 0.3$, $t = 0.4$ and $t = 0.5$. However, for $t = 0.6$, the rate of convergence is no more than $1/2$

With respect to numerics, we described our discretisation methods for approximating TQG and α -TQG solutions and showed that the FEM semi-discrete scheme conserves numerical energy in the TQG case. Using example simulation results, we numerically verified that α -TQG solutions converge to that of the TQG system and attains the order 1 in α convergence rate predicted by Proposition 4.1. Additionally, we derived a dispersion relation for the linearised α -TQG system and showed how α -regularisation could be used to control the development of high wavenumber instabilities. Given that we have convergence of α -TQG solutions to TQG, the linear stability results can be viewed as generalisations of those shown in Holm et al. (2021) for the TQG system.

We now end this section with some open problems.

- Can one construct a global-in-time strong solution (or a global weak solution) of either the TQG model equations (1.1)–(1.5), or the α -TQG model Eqs. (2.12)–(2.14)?
- Is it possible to establish uniform estimates in α for all three blow-up criteria for α -TQG?
- Can one give a Beale–Kato–Majda condition for the blow-up of a strong solution to the TQG in terms of a single unknown? That is, is there a BKM condition in terms of either b or ω (or \mathbf{u}) that does not require a combination of both variables?
- As discussed in Sect. 1.3, an application of the Stochastic Advection by Lie Transport (SALT) approach to the Hamiltonian formulation of the TQG equations results in a stochastic Hamiltonian formulation of TQG equations in (1.6)–(1.8). This stochastic version of TQG preserves an infinite family of integral conserved quantities, as is shown in Holm et al. (2021). The SALT version of TQG represents an outstanding challenge for uncertainty quantification and data assimilation which will surely spur us on to further investigations of these problems. The analysis of the stochastic thermal quasi-geostrophic equation has been covered in the companion paper (Crisan et al. 2022).

Acknowledgements We would like to thank our friends and colleagues for their encouraging comments, especially W. Bauer, F. J. Beron-Vera, B. Chapron, L. Cope, C. J. Cotter, E. Dinvai, O. Lang, E. Memin, A. Radomska - Botelho Moniz, and S. Takao. We would also like to express our sincere gratitude towards the anonymous reviewers.

Author Contributions DH and EL performed the modelling. PM and DC performed the theoretical analysis. WP performed the numerical analysis and the simulations. All authors reviewed the manuscript.

Funding This work was supported by European Research Council (ERC) Synergy grant STUOD-DLV-856408.

Availability of data and material Data sharing is not applicable to this article as no datasets were generated or analysed during the current study. The simulations were performed with the automated system for the solution of partial differential equations using finite elements named *Firedrake*, Rathgeber et al. (2016). The GitHub repository of can be found here <https://github.com/firedrakeproject/firedrake>.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this work. All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

8 Appendix

We present a result in this section, which, along with its proof, can be found in Kato and Lai (1984). Our construction of local solutions relies on this result, and thus, we present it here for the sake of completeness.

We consider the abstract equation

$$\frac{\partial}{\partial t} u + \mathcal{A}(t, u) = 0, \quad t \geq 0, \quad u|_{t=0} = u_0, \quad (8.1)$$

where \mathcal{A} is a nonlinear operator.

Theorem 8.1 *Let $\{V, H, X\}$ be real separable Banach spaces such that*

- *the embeddings $V \hookrightarrow H \hookrightarrow X$ are continuous and dense;*
- *H is a Hilbert space;*
- *There is a continuous, non-degenerate bilinear form (\cdot, \cdot) on $X \times V$ such that $(u, v) = \langle u, v \rangle_H$ for $u \in H$ and $v \in V$.*

Let \mathcal{A} be a (sequentially) weakly continuous map on $[0, T_] \times H$ into X such that*

$$(\mathcal{A}(t, v), v) \geq -\rho(\|v\|_H^2) \quad \text{for } t \in [0, T_*], \quad v \in V, \quad (8.2)$$

where $\rho(r) \geq 0$ is a monotone increasing function of $r \geq 0$. Then for any $u_0 \in H$, there exists $T > 0$, $T \leq T_$, and a solution u of (8.1) in the class*

$$u \in C_w([0, T]; H) \cap C_w^1([0, T]; X). \quad (8.3)$$

Moreover,

$$\|u(t)\|_H^2 \leq r(t), \quad t \in [0, T],$$

where r is a monotone increasing function on $[0, T]$; T and r can be chosen so as to depend only on ρ and $\|u_0\|_H$.

Remark 8.2 (a) If \mathcal{A} maps $[0, T_*] \times V$ into H , we may replace (8.2) with

$$\langle \mathcal{A}(t, v), v \rangle_H \geq -\rho(\|v\|_H^2) \quad \text{for } t \in [0, T_*], \quad v \in V. \quad (8.4)$$

(b) T and r can be determined by solving the scalar differential equation

$$\frac{d}{dt}r = 2\rho(r), \quad r(0) = \|u_0\|_H^2. \quad (8.5)$$

T may be any value such that r exists on $[0, T]$. If the solution to (8.5) is not unique, r should be the maximal solution.

(c) $u(t) \rightarrow u_0$ holds strongly in H as $t \rightarrow 0$. In other words, the solution (8.3) is strongly continuous at the initial time $t = 0$.

References

- Anderson, D.L.T., McCreary, J.P., Jr.: On the role of the Indian Ocean in a coupled ocean-atmosphere model of El Niño and the Southern Oscillation. *J. Atmos. Sci.* **42**(22), 2439–2442 (1985)
- Beale, J.T., Kato, T., Majda, A.: Remarks on the breakdown of smooth solutions for the 3-D Euler equations. *Commun. Math. Phys.* **94**(1), 61–66 (1984)
- Beier, E.: A numerical investigation of the annual variability in the Gulf of California. *J. Phys. Oceanogr.* **27**(5), 615–632 (1997)
- Bernsen, E., Bokhove, O., van der Vegt, J.J.W.: A (dis) continuous finite element model for generalized 2d vorticity dynamics. *J. Comput. Phys.* **211**(2), 719–747 (2006)
- Beron-Vera, F.J.: Nonlinear saturation of thermal instabilities. *Phys. Fluids* **33**(3), 036608 (2021)
- Beron-Vera, F.J.: Multilayer shallow-water model with stratification and shear. *Revista mexicana de física* **67**(3), 351–364 (2021)
- Brenner, S.C., Scott, L.R., Scott, L.R.: *The Mathematical Theory of Finite Element Methods*, vol. 3. Springer, Berlin (2008)
- Brezis H, Gallouet T: Nonlinear Schrödinger evolution equations. Technical report, Wisconsin Univ-Madison Mathematics Research Center, (1979)
- Crisan D, Holm DD, Lang O, Mensah PR, Wei P: Theoretical analysis and numerical approximation for the stochastic thermal quasi-geostrophic model. [arXiv:2207.07457](https://arxiv.org/abs/2207.07457), (2022)
- Dinvey, E.: Well-posedness for a Whitham–Boussinesq system with surface tension. *Math. Phys. Anal. Geom.* **23**(2), 1–27 (2020)
- Foias, C., Holm, D.D., Titi, E.S.: The Navier–Stokes-alpha model of fluid turbulence. *Physica D Nonlinear Phenom.* **152**, 505–519 (2001)
- Foias, C., Holm, D.D., Titi, E.S.: The three dimensional viscous Camassa–Holm equations, and their relation to the Navier–Stokes equations and turbulence theory. *J. Dyn. Differ. Equ.* **14**(1), 1–35 (2002)
- Gibson, T.H., McRae, A.T.T., Cotter, C.J., Mitchell, L., Ham, D.A.: *Compatible Finite Element Methods for Geophysical Flows: Automation and Implementation Using Firedrake*. Springer, Berlin (2019)
- Hesthaven, J.S., Warburton, T.: *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer, Berlin (2007)
- Holm, D.D.: Variational principles for stochastic fluid dynamics. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **471**(2176), 20140963 (2015)
- Holm, D.D., Luesink, E.: Stochastic wave-current interaction in thermal shallow water dynamics. *J. Nonlinear Sci.* **31**(2), 1–56 (2021)
- Holm, D.D., Wingate, B.A.: Baroclinic instabilities of the two-layer quasigeostrophic alpha model. *J. Phys. Oceanogr.* **35**(7), 1287–1296 (2005)
- Holm, D.D., Marsden, J.E., Ratiu, T.S.: The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.* **137**(1), 1–81 (1998a)
- Holm, D.D., Marsden, J.E., Ratiu, T.S.: Euler–Poincaré models of ideal fluids with nonlinear dispersion. *Phys. Rev. Lett.* **80**(19), 4173 (1998b)

- Holm, D.D., Luesink, E., Pan, W.: Stochastic mesoscale circulation dynamics in the thermal ocean. *Phys. Fluids* **33**(4), 046603 (2021)
- Kato, T.: Liapunov functions and monotonicity in the Navier–Stokes equation. In: *Functional-analytic Methods for Partial Differential Equations*, pp. 53–63. Springer (1990)
- Kato, T., Lai, C.Y.: Nonlinear evolution equations and the Euler flow. *J. Funct. Anal.* **56**(1), 15–28 (1984)
- Kato, T., Ponce, G.: Commutator estimates and the Euler and Navier–Stokes equations. *Commun. Pure Appl. Math.* **41**(7), 891–907 (1988)
- Klainerman, S., Majda, A.: Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids. *Commun. Pure Appl. Math.* **34**(4), 481–524 (1981)
- Marsden, J.E., Shkoller, S.: The anisotropic Lagrangian averaged Euler and Navier–Stokes equations. *Arch. Ration. Mech. Anal.* **166**(1), 27–46 (2003)
- McCreary, J.P., Jr., Zhang, S., Shetye, S.R.: Coastal circulations driven by river outflow in a variable-density 11/2-layer model. *J. Geophys. Res. Oceans* **102**(C7), 15535–15554 (1997)
- McWilliams, J.C.: A survey of submesoscale currents. *Geosci. Lett.* **6**(1), 1–15 (2019)
- O’Brien, J.J., Reid, R.O.: The non-linear response of a two-layer, baroclinic ocean to a stationary, axially-symmetric hurricane: part I. Upwelling induced by momentum transfer. *J. Atmos. Sci.* **24**(2), 197–207 (1967)
- Rathgeber, F., Ham, D.A., Mitchell, L., Lange, M., Luporini, F., McRae, A.T.T., Bercea, G.-T., Markall, G.R., Kelly, P.H.J.: Firedrake: automating the finite element method by composing abstractions. *ACM Trans. Math. Softw. (TOMS)* **43**(3), 1–27 (2016)
- Ripa, P.: Conservation laws for primitive equations models with inhomogeneous layers. *Geophys. Astrophys. Fluid Dyn.* **70**(1–4), 85–111 (1993)
- Ripa, P.: On improving a one-layer ocean model with thermodynamics. *J. Fluid Mech.* **303**, 169–201 (1995)
- Ripa, P.: Low frequency approximation of a vertically averaged ocean model with thermodynamics. *Revista Mexicana de Física* **42**(1), 117–135 (1996)
- Ripa, P.: On the validity of layered models of ocean dynamics and thermodynamics with reduced vertical resolution. *Dyn. Atmos. Oceans* **29**(1), 1–40 (1999)
- Schopf, P.S., Cane, M.A.: On equatorial dynamics, mixed layer physics and sea surface temperature. *J. Phys. Oceanogr.* **13**(6), 917–935 (1983)
- Volkov, D.L., Kubryakov, A.A., Lumpkin, R.: Formation and variability of the lofoten basin vortex in a high-resolution ocean model. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **105**, 142–157 (2015)
- Wang, C., Zhang, Z.: Global well-posedness for the 2-D Boussinesq system with the temperature-dependent viscosity and thermal diffusivity. *Adv. Math.* **228**(1), 43–62 (2011)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.