University of Reading

Exponential ergodicity for a stochastic two-layer quasi-geostrophic model

Article

Accepted Version

Carigi, G. ORCID: https://orcid.org/0000-0001-7611-8230,
Bröcker, J. and Kuna, T. (2023) Exponential ergodicity for a stochastic two-layer quasi-geostrophic model. Stochastics and Dynamics, 23 (2). 2350011. ISSN 1793-6799 doi: https://doi.org/10.1142/S0219493723500119 Available at https://centaur.reading.ac.uk/109663/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1142/S0219493723500119

Publisher: World Scientific Publishing

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

arXiv:2201.09823v1 [math.PR] 24 Jan 2022

Exponential ergodicity for a stochastic two-layer quasi–geostrophic model

Giulia Carigi*† Jochen Bröcker*†‡ Tobias Kuna*†§

January 25, 2022

Abstract

Ergodic properties of a stochastic medium complexity model for atmosphere and ocean dynamics are analysed. More specifically, a two–layer quasi–geostrophic model for geophysical flows is studied, with the upper layer being perturbed by additive noise. This model is popular in the geosciences, for instance to study the effects of a stochastic wind forcing on the ocean. A rigorous mathematical analysis however meets with the challenge that in the model under study, the noise configuration is spatially degenerate as the stochastic forcing acts only on the top layer. Exponential convergence of solutions laws to the invariant measure is established, implying a spectral gap of the associated Markov semigroup on a space of Hölder continuous functions. The approach provides a general framework for generalised coupling techniques suitable for applications to dissipative SPDE's. In case of the two–layer quasi–geostrophic model, the results require the second layer to obey a certain passivity condition.

Keywords: generalized coupling; SPDEs; Stochastic geophysical flow models.

AMS Subject Classification: *Primary:* 37L40, 76U60 *Secondary:* 60H15, 86A08

1 Introduction

In this work we study the long time average behaviour of a stochastic version of an important model for large-scale atmosphere and ocean dynamics, the two–layer quasi–geostrophic (2LQG) model with a forcing on the top layer, to account, for example, for the wind forcing on the upper ocean, composed of

*Department of Mathematics and Statistics, University of Reading, Reading, UK
†Centre for the Mathematics of Planet Earth, University of Reading, Reading, UK
‡Department of Meteorology, University of Reading, Reading, UK
§Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila, 67100 L'Aquila, Italy

a deterministic and a stochastic part, which is white in time and coloured in space. We will provide conditions for the model to be exponentially ergodic, which means that the transition probabilities converge to the unique stationary probability distribution with an exponential rate, and that temporal averages of an observable converge to averages of the observable with respect to that stationary probability distribution. Indeed, showing exponential ergodicity for the stochastic 2LQG model and, in particular ergodicity, justifies that the long term average behaviour of the ocean dynamics at the mid–latitudes under stochastic wind stress can be studied using the unique probability distribution invariant under the dynamics. In general ergodicity is often a tacit assumption underlying applications, as it allows to draw conclusions on the statistics of the system from its time series.

In mathematical terms, for a system modeled by a stochastic partial differential equation (SPDE), as in this paper, exponential ergodicity means the following: Consider the Markov semigroup and its transition probabilities $\{P_t, t \geq 0\}$ associated to it. When μ is an invariant measure with respect to the transition probabilities, we speak of exponential ergodicity if the transition probabilities $\{P_t, t \geq 0\}$ converge exponentially in time to an invariant measure μ, independent of the initial data. This implies in particular that there exists exactly one invariant measure which has thus to be ergodic in the usual sense.

The quasi–geostrophic (QG) model, already present in the literature from the late 1930s, was systematically derived by Charney in 1948 in [9]. Used in early operational numerical weather forecasts, the QG model is still used extensively in research as it is strikes a balance between simplicity of formulation versus spectrum of the phenomena it can reproduce. The QG model is an approximation of three dimensional Navier–Stokes equations on a rotating coordinate frame which best models the large-scale features (e.g. 1000 km for the atmosphere and 100 km for the ocean) of the atmosphere or the ocean at mid latitudes. Quasi–geostrophic models with several layers in particular are able to represent density stratification and provide insights into, for instance, atmosphere–ocean coupling and baroclinic instabilities. This type of instabilities is extremely common in both the atmosphere and ocean and is at the origin of large scale weather phenomena, for instance mid–latitude cyclones. The two–layer quasi–geostrophic (2LQG) model is one of the simplest models where the baroclinic instabilities arise.

The effect of a stochastic wind forcing on QG models has been a topic of research in meteorology and oceanography for at least thirty years, both in the single layer case e.g. [20, 29, 31] and the multi–layer case, see e.g. [2, 10, 13, 30]. In the mathematical literature the stochastic single–layer QG model, with either additive or multiplicative noise, has definitely received more attention (see e.g. [4, 14, 15, 34]) than its multi–layer version. In fact, we can expect results achieved for the single layer to extend to the multi–layer case in situations where the random terms appear in all layers. Less studied, though, is the action of a stochastic forcing acting only on one of the layers and its consequent effects on the other layers and the whole dynamics.

To the best of the authors' knowledge, the only reference for a mathemati-

cal investigation of a two–layer quasi–geostrophic model with a forcing on the top layer, prior to this work, is [11]. There the authors studied the long time dynamics of the model using the method of determining functionals for random dynamical systems. This method gives a way to parametrise the system's global attractor by means of a finite number of functionals. Furthermore, under some conditions on the parameters of the system, it is shown that functionals depending only on the top layer suffice to describe the attractor. However, this approach does not give information on the statistics of the model which will be the focus here.

With regards to exponential ergodicity, Harris' theorem provides conditions under which exponential convergence of transition probabilities holds in the total variation norm. A major obstacle to applying this theorem in the context of SPDEs is that the transition probabilities may very easily be singular for different initial conditions. In [24, Section 4] a new framework is introduced to retrieve a version of Harris' theorem in the infinite dimensional context, which gives exponential rate of convergence in a Wasserstein semimetric, rather than in total variation. The uniqueness and hence ergodicity of the invariant measure follow as well, and in most cases of interest also its existence.

However the assumptions required to apply this generalised Harris theorem may be challenging to show in practice. Recently [6] provided a set of conditions which gives exponential and sub-exponential rate of convergence to the invariant measure in the Wasserstein semimetric, improving [5] and [25]. The approach in [6] relies on the *generalised coupling* method. The main idea of this method is to add a control to the stochastic forcing to estimate the distance between solutions with different initial data, or more precisely estimate the Wasserstein semimetric between their laws. This approach was introduced in the early 2000's (see e.g. [27, 21, 24]) as *asymptotic* coupling method, but in this case the added control ensures that the solutions with different initial data synchronise, namely have the same law asymptotically in time, at least with positive probability. The term *generalised coupling* was introduced in [25].

The results in [6] are fairly general and apply to a wide variety of infinite dimensional systems. In the present paper, we present a methodology inspired by the results in [6] but with a view on applications to stochastic dissipative SPDE's that describe atmosphere and ocean dynamics. For such SPDE's we provide conditions sufficient for the results of [6] (in particular the generalised Harris' theorem) to apply. Our conditions are straightforward to check and easy to interpret mathematically and physically. For the convenience of the reader, we provide proofs for those parts of [6] where our line of reasoning deviates from [6].

We will apply this technique to the stochastic two–layer quasi–geostrophic model to show exponential convergence of transition probabilities. Since the noise is acting only on one of the layers, the control will have to appear only in that layer but still be able to stabilise the lower layer. We will be able to find such a control by also imposing a condition involving the bottom friction to ensure the result. We may think of the imposed parameter condition as requiring the bottom layer (the one without noise) to be sufficiently dissipative so as to

be determined by the top one. This provides information on the stability of the system for large times, in the sense that not only is there a unique invariant measure but the system approaches it exponentially fast in time, provided the bottom friction is large enough.

The results presented in this work are applicable (with appropriate modifications) to models with more than two layers. The condition on the passivity of the bottom layer would have to apply to the additional layers without stochastic forcing. Although the passivity condition might still be considered realistic in two layers with the second accounting for the lower ocean due to the friction with bottom of the ocean for example, instead for multiple layers a strong friction on all layers is harder to justify from a physical point of view, especially since the density difference between the layers (which would basically determine the magnitude of this friction) is not so sharp. How to remove this restriction will be subject to future research.

An important immediate consequence of the exponential convergence of the transition probabilities is the *spectral gap* for the Markov semigroup on a suitable space of functions. In our case we clarify that these functions are locally Hölder continuous. The spectral gap is a crucial ingredient to prove linear response, which investigates how the invariant measure depends on the parameters of the system (see e.g. [22]). Linear response is a rigorous mathematical framework in which climate change triggered by changes in external parameters or forcings can be investigated (see e.g. [1, 18]). Studying linear response for the 2LQG model was one of the motivations behind establishing the spectral gap, and we will consider linear response in a forthcoming paper [8].

Overview over the results The stochastic two–layer quasi–geostrophic model, to be presented in Section 2 in more detail, describes two layers of fluid on top of one another with certain heights and densities. Both Coriolis effects and eddy viscosity is taken into account. Forcing acts only on the top layer and has a non-trivial stochastic part which accounts for example for the effect of the wind on the upper ocean (with the second layer then representing the deeper ocean).

To describe the model mathematically, let $\mathcal{D}$ be a squared domain $\mathcal{D} = [0, L] \times [0, L] \subset \mathbb{R}^2$ (e.g. L is of the order 10^6m for the atmosphere and 10^5m for the ocean), and consider the equations

$$\begin{aligned} dq_1 + J(\psi_1, q_1 + \beta y)\, dt &= \big(\nu\Delta^2\psi_1 \ + f\big)\, dt + dW \\ \partial_t q_2 + J(\psi_2, q_2 + \beta y) &= \nu\Delta^2\psi_2 - r\Delta\psi_2, \end{aligned} \tag{1}$$

where $\mathbf{x} = (x, y) \in \mathcal{D}$, $\boldsymbol{\psi}(t, \mathbf{x}) = (\psi_1(t, \mathbf{x}), \psi_2(t, \mathbf{x}))^t$ is the streamfunction of the fluid, and $\mathbf{q}(t, \mathbf{x}) = (q_1(t, \mathbf{x}), q_2(t, \mathbf{x}))^t$ is the so-called quasi–geostrophic potential vorticity. Vorticity and streamfunction are related through $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$ with an elliptic operator $\tilde{A}$ which also includes physical constants and parameters of the model (see Eq. (13)). Moreover, J is the Jacobian operator $J(a, b) = \nabla^\perp a \cdot \nabla b$, while dW represents noise which is white in time but coloured in space, with trace class covariance operator Q. Finally, the model includes a (time-independent) deterministic forcing on the top layer $f = f(\mathbf{x})$.

Conditions ensuring the well–posedness of this model, along with further auxiliary results, are presented in Theorem 2.4.

In Section 3 we consider a generic dissipative SPDE on a Hilbert space $(\mathcal{H}, |\cdot|)$ with associated Markov semigroup $\mathcal{P}_t$ which is Feller and satisfies Assumption A (see below). Theorem 3.7 demonstrates that the transition probabilities are a strict contraction in a Wasserstein semimetric associated with a suitable semimetric $\tilde{d}$ on $(\mathcal{H}, |\cdot|)$. This implies, in particular, the existence of a spectral gap (see Corollary 3.5) as well as the generalised Harris' Theorem 3.3. The core Assumption A of Theorem 3.7 contains a priori type energy estimates on the solution of the SPDE (**A2**), requirements on the control (**A1**, stabilisation of the dynamics, and **A3**, regularity), and the existence of an appropriate Lyapunov function (**A4**).

In Section 4, the results from Section 3 are applied to the stochastic 2LQG model in Equation (1), giving us the following result (recall that r determines the bottom friction, see 2nd line of Eq. 1):

Theorem 4.1 (See Sec. 4 for precise statements)**.** *There exists r_0 (depending on ν, Q and the forcing f from Eq. 1) so that if $r > r_0$, and* $\operatorname{range} Q$ *is large enough (depending on r in a sense to be made precise), then there exists a unique invariant measure μ_* as well as $t > 0$ and $\rho < 1$ such that*

$$W_{\tilde{d}}\left(P_t(\mathbf{q}_0, \cdot), P_t(\tilde{\mathbf{q}}_0, \cdot)\right) \leq \rho\, \tilde{d}(\mathbf{q}_0, \tilde{\mathbf{q}}_0) \tag{2}$$

for all $\mathbf{q}_0, \tilde{\mathbf{q}}_0 \in \mathcal{H}$.

The proof proceeds by showing that Assumption A is satisfied. In Remark 4.5 it is demonstrated that Theorem 4.1 also holds for any given r and Q, provided the viscosity ν is sufficiently large, see (77). The semimetric $\tilde{d}$ appearing in Theorem 3.7 is qualitatively of the form $\tilde{d}(x, y) \cong |x - y|^\alpha$ for small $x, y \in \mathcal{H}$, with $\alpha < \frac{1}{2}$. Hence our spectral gap result Corollary 3.5 refers to functions which are, roughly speaking, of Hölder type.

Acknowledgments The work presented here would have been impossible were it not for fruitful discussions with a number of colleagues. In particular, we are very grateful to Benedetta Ferrario, Franco Flandoli, Valerio Lucarini, and Jeroen Wouters for criticisms, comments, suggestions, and encouragement. GC's work was funded by the Centre for Doctoral Training in Mathematics of Planet Earth, UK EPSRC funded (grant EP/L016613/1), by the LMS Early Career Fellowship (grant ECF1920-48), and by the UK EPSRC grant EP/W522375/1. Furthermore, GC would like to thank the Institute Henri Poincaré for supporting the participation to the thematic trimester on *The Mathematics of Climate and the Environment* in the Autumn of 2019.

2 The stochastic 2LQG model

The two–layer quasi–geostrophic model has been described as "perhaps the most widely used set of equations for theoretical studies of atmosphere and

ocean" [32]. From a mathematical point of view the 2LQG model is a system of several 2D Navier–Stokes equations in vorticity formulation coupled to each other. In this section we lay down its precise mathematical formulation and the main notations following closely the set up described in [3] and [11].

2.1 Mathematical setup and notation

The 2LQG equations are modelling two layers of fluid one on top of each other with mean height h_1 for the top layer and h_2 for the bottom one, and with density respectively ρ_1 and ρ_2 with $\rho_1 < \rho_2$ We consider the so-called β-plane approximation to the Coriolis effect (see [32, Section 2.3.2]); this accounts for the fact that the vertical component of the rotation changes with the latitude y by writing the Coriolis parameter as $f_c(y) = f_0 + \beta y$, with f_0 and β assigned positive constants. We also take into account the effect of the eddy viscosity on both layers, and of the the bottom friction on the second layer. We assume that the forcing acts only on the top layer and has a non-trivial stochastic part which accounts for example for the effect of the wind on the upper ocean.

Let $\mathcal{D}$ be a squared domain $\mathcal{D} = [0, L] \times [0, L] \subset \mathbb{R}^2$. Consider the following equations

$$\begin{aligned} dq_1 + J(\psi_1, q_1 + \beta y)\, dt &= \left(\nu\Delta^2\psi_1 \ + f\right) dt + dW \\ \partial_t q_2 + J(\psi_2, q_2 + \beta y) &= \nu\Delta^2\psi_2 - r\Delta\psi_2 \end{aligned} \tag{3}$$

where $\mathbf{x} = (x, y) \in \mathcal{D}$, $\boldsymbol{\psi}(t, \mathbf{x}) = (\psi_1(t, \mathbf{x}), \psi_2(t, \mathbf{x}))^t$ is the streamfunction of the fluid, and $\mathbf{q}(t, \mathbf{x}) = (q_1(t, \mathbf{x}), q_2(t, \mathbf{x}))^t$ is the so-called quasi–geostrophic potential vorticity. Vorticity and streamfunction are related through

$$\begin{aligned} q_1 &= \Delta\psi_1 + F_1(\psi_2 - \psi_1) \\ q_2 &= \Delta\psi_2 + F_2(\psi_1 - \psi_2) \end{aligned} \tag{4}$$

where F_1, F_2 are positive constants. Moreover, J is the Jacobian operator $J(a, b) = \nabla^\perp a \cdot \nabla b$. W is a random term, white in time and colored in space, more precisely, a so-called Q-Wiener process, which accounts for the stochastic part of the forcing (more details in Sec. 2.2). Furthermore we assume periodic boundary conditions for $\boldsymbol{\psi}$ in both directions with period L and we impose that

$$\int_{\mathcal{D}} \boldsymbol{\psi}(t, \mathbf{x})\, d\mathbf{x} = 0 \quad \text{for all } \ t \geq 0. \tag{5}$$

The model includes a deterministic forcing on the top layer $f = f(\mathbf{x})$ (time-independent) as well with zero spatial averages, i.e.

$$\int_{\mathcal{D}} f(\mathbf{x})\, d\mathbf{x} = 0.$$

The constants F_1, F_2 are related to physical constants by

$$F_i := \frac{f_0^2}{g' h_i}, \tag{6}$$

with g' the reduced gravity, $g' = g(\rho_2 - \rho_1)/\rho_0$ where $\rho_0 = (\rho_1 + \rho_2)/2$ is the characteristic value for the density, and g is the gravitational acceleration. Further, denote by

$$h_1 F_1 = h_2 F_2 = \frac{f_0^2}{g'} =: p. \tag{7}$$

The model (3) includes dissipation generated by the eddy viscosity on both layers modeled by the terms $\nu\Delta^2\psi_i$ and by the friction with the bottom modeled by $r\Delta\psi_2$. We can write (3) in vectorial formulation introducing

$$B(\boldsymbol{\psi}, \boldsymbol{\xi}) = \begin{pmatrix} J(\psi_1, \Delta\xi_1) + F_1 J(\psi_1, \xi_2) \\ J(\psi_2, \Delta\xi_2) + F_2 J(\psi_2, \xi_1) \end{pmatrix}. \tag{8}$$

Using the fact that $J(\psi, \psi) = 0$, we can write more compactly (3) as

$$d\mathbf{q} + (B(\boldsymbol{\psi}, \boldsymbol{\psi}) + \beta\partial_x \boldsymbol{\psi})\, dt = \nu\Delta^2\boldsymbol{\psi}\, dt + \begin{pmatrix} f \\ -r\Delta\psi_1 \end{pmatrix} dt + d\mathbf{W} \tag{9}$$

where $\mathbf{W} = (W, 0)^t$, and $\Delta\boldsymbol{\psi} = (\Delta\psi_1, \Delta\psi_2)^t$. Moreover, we can express (4) vectorial as well

$$\mathbf{q} = (\Delta + M)\boldsymbol{\psi} \quad \text{with } M = \begin{pmatrix} -F_1 & F_1 \\ F_2 & -F_2 \end{pmatrix}. \tag{10}$$

Next we introduce the notations for the mathematical setup we want to consider for the two–layer quasi–geostrophic. Let $(L^2(\mathcal{D}), \|\cdot\|_0)$, $(H^k(\mathcal{D}), \|\cdot\|_k)$, $k \in \mathbb{R}$ be the standard Sobolev spaces of L-periodic functions satisfying (5). Denote by $(\cdot,\cdot)_k$ the associated scalar product. We also introduce appropriate norms on the product spaces to deal with our coupled system. Given $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$ elements of $H^k \times H^k$, $k > 0$ or $L^2 \times L^2$ for $k = 0$, define

$$\begin{aligned} (\boldsymbol{\psi}, \boldsymbol{\xi})_k &:= h_1(\psi_1, \xi_1)_k + h_2(\psi_2, \xi_2)_k \\ \|\boldsymbol{\psi}\|_k^2 &:= h_1\|\psi_1\|_k^2 + h_2\|\psi_2\|_k^2. \end{aligned} \tag{11}$$

Then we define

$$\begin{aligned} \mathbf{L}^2 &= \left\{\boldsymbol{\psi} \in L^2 \times L^2 \,:\, \|\boldsymbol{\psi}\|_0^2 < \infty\right\} \\ \mathbf{H}^k &= \left\{\boldsymbol{\psi} \in H^k \times H^k \,:\, \|\boldsymbol{\psi}\|_k^2 < \infty\right\}, \quad k > 0 \end{aligned}$$

and we denote with $\mathbf{H}^{-k}$ the dual space of $\mathbf{H}^k$, $k > 0$.

The Poincaré inequality in $\mathbf{H}^k$ reads as

$$\|\boldsymbol{\psi}\|_k \le \lambda_1^{-1/2}\|\boldsymbol{\psi}\|_{k+1}, \tag{12}$$

where λ_1 is the smallest eigenvalue of the operator $-\Delta$.

Define the operator $\tilde{A} : \mathbf{H}^{k+2} \to \mathbf{H}^k$, $k \in \mathbb{R}$, connection of the streamfunction with the quasi–geostrophic potential vorticity

$$\tilde{A}\boldsymbol{\psi} = -(\Delta + M)\boldsymbol{\psi}, \quad \boldsymbol{\psi} \in \mathbf{H}^{k+2}. \tag{13}$$

It is easy to see that $\tilde{A}$ is an unbounded non–negative self–adjoint operator in $\mathbf{H}^k$ with respect to the weighted scalar product $(\cdot,\cdot)_k$, indeed

$$(\tilde{A}\boldsymbol{\psi},\boldsymbol{\psi})_k = -(\Delta\boldsymbol{\psi},\boldsymbol{\psi})_k - (M\boldsymbol{\psi},\boldsymbol{\psi})_k = \|\boldsymbol{\psi}\|_{k+1}^2 + p|\psi_1-\psi_2|_k^2, \tag{14}$$

which, due to (5), has an inverse which is bounded as function $\mathbf{H}^k \to \mathbf{H}^{k+2}$. Then for each $\mathbf{q} \in \mathbf{H}^k$ there exists $\boldsymbol{\psi} \in \mathbf{H}^{k+2}$ such that $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$. It is interesting to note that if the upper layer is more dense than the bottom, namely $\rho_1 \geq \rho_2$, then F_1, F_2 would change sign and $\tilde{A} = -(\Delta - M)$. Therefore $\tilde{A}$ would not be non–negative and one could not develop a consistent theory for such equations, reflecting the physical impossibility of this setup.

Remark 2.1. Since the $\mathbf{L}^2$ and $\mathbf{H}^1$ norms and the $\mathbf{L}^2$ scalar product are the most used throughout this work, for the sake of simplifying notation we denote them as follows

$$\begin{aligned} |\psi| := \|\psi\|_0 \quad &\text{and} \quad \|\psi\| := \|\psi\|_1 \\ |\boldsymbol{\psi}| = h_1|\psi_1| + h_2|\psi_2| := \|\boldsymbol{\psi}\|_0 \quad &\text{and} \quad \|\boldsymbol{\psi}\| = h_1\|\psi_1\| + h_2\|\psi_2\| := \|\boldsymbol{\psi}\|_1 \\ (\psi,\xi) := (\psi,\xi)_0 \quad &\text{and} \quad (\boldsymbol{\psi},\boldsymbol{\xi}) = h_1(\psi_1,\xi_1) + h_2(\psi_2,\xi_2) \end{aligned}$$

Finally, we introduce two new norms on the level of the potential vorticities. For $\mathbf{q} \in \mathbf{H}^{-1}$ there exists $\boldsymbol{\psi} \in \mathbf{H}^1$ such that $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$, and we can define the norm on $\mathbf{H}^{-1}$

$$|||\mathbf{q}|||_{-1}^2 := \|\boldsymbol{\psi}\|^2 + p|\psi_1-\psi_2|^2$$

and, for $\mathbf{q} \in \mathbf{L}^2$ with $\boldsymbol{\psi} \in \mathbf{H}^2$ define the norm on $\mathbf{L}^2$

$$|||\mathbf{q}|||_0^2 := |\Delta\boldsymbol{\psi}|^2 + p\|\psi_1-\psi_2\|^2.$$

Note that by Poincaré inequality one has

$$\begin{aligned} |||\mathbf{q}(t)|||_{-1}^2 &= \|\boldsymbol{\psi}\|^2 + p|\psi_1-\psi_2|^2 \\ &\leq \lambda_1^{-1}\left(|\Delta\boldsymbol{\psi}|^2 + p\|\psi_1-\psi_2\|^2\right) = \lambda_1^{-1}|||\mathbf{q}(t)|||_0^2. \end{aligned}$$

Furthermore, these norms are equivalent to $\|\cdot\|_{-1}$ and $\|\cdot\|_0$ respectively and have a series of useful properties, which we will show now and exploit later on.

Lemma 2.2. *Consider* $\mathbf{q} \in \mathbf{H}^{-1}$ *and* $\boldsymbol{\psi} \in \mathbf{H}^1$ *such that* $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$. *Then the following relations hold:*

$$-(\mathbf{q},\boldsymbol{\psi}) = |||\mathbf{q}|||_{-1}^2 \tag{15}$$

$$\|\boldsymbol{\psi}\|^2 \leq |||\mathbf{q}|||_{-1}^2 \leq a_0\|\boldsymbol{\psi}\|^2 \tag{16}$$

for some $a_0 > 0$. *For* $\mathbf{q} \in \mathbf{L}^2$ *and* $\boldsymbol{\psi} \in \mathbf{H}^2$ *such that* $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$, *we have:*

$$(\mathbf{q},\Delta\boldsymbol{\psi}) = |||\mathbf{q}|||_0^2 \tag{17}$$

$$|\Delta\boldsymbol{\psi}|^2 = |\Delta\boldsymbol{\psi}|^2 \leq |||\mathbf{q}|||_0^2 \leq a_0|\Delta\boldsymbol{\psi}|^2. \tag{18}$$

Proof. Equality (15) is a direct calculation, cf. also (14). In a similar way we can show (17). By definition of $\mathbf{q}$ we have

$$(\mathbf{q}, \Delta\boldsymbol{\psi}) = (\Delta\boldsymbol{\psi}, \Delta\boldsymbol{\psi}) + (M\boldsymbol{\psi}, \Delta\boldsymbol{\psi}).$$

Then by definition of M (10), relation (7) and Green's theorem

$$\begin{aligned}(M\boldsymbol{\psi}, \Delta\boldsymbol{\psi}) &= -h_1 F_1(\psi_1 - \psi_2, \Delta\psi_1) + h_2 F_2(\psi_1 - \psi_2, \Delta\psi_2) \\ &= p((-\Delta)^{1/2}(\psi_1 - \psi_2), (-\Delta)^{1/2}\psi_1) - p((-\Delta)^{1/2}(\psi_1 - \psi_2), (-\Delta)^{1/2}\psi_2).\end{aligned}$$

Therefore we have

$$(\mathbf{q}, \Delta\boldsymbol{\psi}) = |\Delta\boldsymbol{\psi}|^2 + p\|\psi_1 - \psi_2\|^2 = |||\mathbf{q}|||_0^2.$$

Moving on to (16), the lower bound follows from (14) and the upper bound is a consequence of the Poincaré inequality (12) and the parallelogram law. Indeed we have

$$\begin{aligned}p|\psi_1 - \psi_2|^2 &\leq p\lambda_1^{-1}\|\psi_1 - \psi_2\|^2 \\ &\leq \frac{2p}{\lambda_1 \min(h_1, h_2)}(h_1\|\psi_1\|^2 + h_2\|\psi_2\|^2) \\ &= 2\lambda_1^{-1}\max(F_1, F_2)\|\boldsymbol{\psi}\|^2.\end{aligned}$$

Setting a_0 to be $1 + 2\lambda_1^{-1}\max(F_1, F_2)$, we see that (17) holds.

With similar arguments (18) can be shown as well.

□

Table 1 contains a summary of the spaces and relative norms used throughout this work.

Space	Norm
$\mathbf{H}^k$	$\|\boldsymbol{\psi}\|_k^2 = h_1\|\psi_1\|_k^2 + h_2\|\psi_2\|_k^2$
$\mathbf{L}^2 = \mathbf{H}^0$	$\|\boldsymbol{\psi}\|^2 = h_1\|\psi_1\|^2 + h_2\|\psi_2\|^2$
$\mathbf{H}^1$	$\|\psi\|^2 = h_1\|\boldsymbol{\psi}_1\|^2 + h_2\|\psi_2\|^2$
$\mathbf{L}^2$	$\|\|\mathbf{q}\|\|_0^2 = \|\boldsymbol{\psi}\|_2^2 + p\|\psi_1 - \psi_2\|^2$
$\mathbf{H}^{-1}$	$\|\|\mathbf{q}\|\|_{-1}^2 = \|\boldsymbol{\psi}\|^2 + p\|\psi_1 - \psi_2\|^2$

Table 1: Notations for the two–layer quasi–geostrophic model. Rows 1–3 will be mainly used for the streamfunctions $\boldsymbol{\psi}$. Rows 4,5 will be mainly used for the potential vorticities $\mathbf{q}$, i.e. functions in the range of the operator $-\tilde{A}$. Note that $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$ in the second block.

Finally, standard bounds on the Jacobian see for example [11, Lemma 3.1], yield the following bound for the bilinearity B:

Lemma 2.3 ([7]). *Let B be the bilinear operator defined in* (8)*, then for* $\boldsymbol{\psi}, \boldsymbol{\xi}, \boldsymbol{\phi} \in \mathbf{H}^2$

$$(B(\boldsymbol{\psi}, \boldsymbol{\xi}), \boldsymbol{\phi}) = -(B(\boldsymbol{\phi}, \boldsymbol{\xi}), \boldsymbol{\psi}), \tag{19}$$

$$(B(\boldsymbol{\psi}, \boldsymbol{\xi}), \boldsymbol{\psi}) = 0. \tag{20}$$

Moreover, for $\boldsymbol{\psi}, \boldsymbol{\xi} \in \mathbf{H}^2$*, there exists positive constant* k_0 *such that*

$$|(B(\boldsymbol{\psi}, \boldsymbol{\psi}), \boldsymbol{\xi})| \leq k_0 \|\boldsymbol{\psi}\| |\Delta \boldsymbol{\psi}| |\Delta \boldsymbol{\xi}|. \tag{21}$$

2.2 The stochastic forcing

On the first layer we consider a forcing with a stochastic component which is white in time and colored in space. More precisely, consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let W be a L^2–valued Wiener process on it with covariance operator Q. We assume Q to be trace class in L^2, namely given $\{e_k\}_{k \in \mathbb{N}}$, a complete orthonormal basis of L^2

$$\operatorname{Tr} Q := \sum_{k=1}^{\infty} \langle Q e_k, e_k \rangle < \infty.$$

We also assume that Q and $-\Delta$ commute, meaning that the $\{e_k\}_{k \in \mathbb{N}}$ can be taken as the eigenvectors of $-\Delta$. Note that all the e_k have to fulfill (5) and hence $\int W(t)(\mathbf{x}) d\mathbf{x} = 0$ as well, but this follows from the fact that we defined the spaces L^2 and H^k so that (5) holds.

For later use we note the following facts. Let $\mathcal{H}$ be a general separable Hilbert space and W a Wiener process with values in $\mathcal{H}$ and trace–class covariance operator Q. It can be shown that $W(t)$ can be written as a sum of real valued Wiener processes, namely

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\sigma_k} \beta_k(t) e_k,$$

where σ_k are eigenvalues of Q with $\{e_k\}_{k \in \mathbb{N}}$ a corresponding orthonormal system of eigenfunctions, and $\beta_k(t)$, $k \in \mathbb{N}$, are independent real valued Brownian motions on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, we define for each $n \in \mathbb{N}$ the orthogonal projection $\Pi_n : \mathcal{H} \to \operatorname{span}\{e_1, \ldots, e_n\}$. Note that (see e.g. [12, Section 4.2.2]) $\Pi_n W$ and $(1 - \Pi_n) W$ are independent Wiener processes with covariance matrices

$$Q_n = \sum_{k=0}^{n} \sigma_k e_k \otimes e_k \quad \text{resp.} \quad Q^n = \sum_{k=n+1}^{\infty} \sigma_k e_k \otimes e_k. \tag{22}$$

Furthermore, let $\tilde{W}_n$ be a n-dimensional Brownian motion, equal in law to $\Pi_n W$. Then, as all β_k are mutually independent, $\tilde{W}(t) := \tilde{W}_n + (1 - \Pi_n) W$ is a Q-Wiener process equal in law to W.

2.3 Solutions of the stochastic 2LQG model

The deterministic version of (9) has been shown to be well–posed in [3], while, for the stochastic model, the following result holds:

Theorem 2.4 ([7]). *Consider the system* (9) *with initial condition* $\mathbf{q}_0 \in \mathbf{H}^{-1}$*, deterministic forcing* $f \in H^{-2}$ *and the covariance operator* Q *trace class in* L^2*. Then there exists a pathwise unique weak solution, namely for a.a.* $\omega \in \Omega$ *there exists a unique* $\mathbf{q}$ *such that*

$$\mathbf{q} \in C([0,T];\mathbf{H}^{-1}) \cap L^2(0,T;\mathbf{L}^2)$$

and that satisfies (9) *in the integral sense:*

$$(\mathbf{q}(t),\varphi) + \int_0^t (B(\boldsymbol{\psi},\boldsymbol{\psi}),\varphi) + \beta\,(\partial_x\boldsymbol{\psi},\varphi)\;ds = (\mathbf{q}_0,\varphi) + \nu\int_0^t (\Delta\boldsymbol{\psi},\Delta\varphi)\;ds$$
$$+ \int_0^t h_1(f,\varphi_1) - rh_2(\Delta\psi_2,\varphi_2)\;ds + h_1\,(W(t),\varphi_1)$$

for all $\varphi = (\varphi_1,\varphi_2)^t \in \mathbf{H}^2$*,* $t \in [0,T]$*. Furthermore,* $\mathbf{q}(t,\mathbf{q}_0)$ *is continuous with respect to the initial condition* $\mathbf{q}_0$ *as a function in* $\mathbf{H}^{-1}$*.*

The well–posedness of Equation (9) was shown in [7], using the same approach as in [17] for Navier–Stokes. The main idea is to reformulate Equation (9) into an equation for $\tilde{\mathbf{q}} := \mathbf{q} - (\eta,\,0)^t$, where η is an auxiliary Ornstein-Uhlenbeck process, namely, a solution η of

$$d\eta - \alpha\Delta\eta\,dt = dW$$

for an appropriate value of the constant $\alpha > 0$. Since the covariance matrix Q is trace class in L^2, by the classic theory of SDEs we know η has a continuous version with values in L^2. As $\tilde{\mathbf{q}}$ now satisfies a deterministic equation with random coefficients, existence and uniqueness can be established with PDE techniques as done for the multi–layer QG model in [3]. In [7] it is shown that there exists a unique $\tilde{\mathbf{q}} \in C([0,T];\mathbf{H}^{-1}) \cap L^2(0,T;\mathbf{L}^2)$ so that $\tilde{\mathbf{q}} + \eta$ is solution of (9). This ensures the well-posedness of (9) as given in Theorem 2.4. Furthermore the solution can be shown to be continuous with respect to the driving noise, meaning that for any $t \geq 0$ and initial condition $\mathbf{q}_0$ there exists a continuous function

$$\Phi_t^0 : C([0,t];L^2) \to \mathbf{H}^{-1}$$

such that $\mathbf{q}(t,\mathbf{q}_0) = \Phi_t^0\left(\{W(s)\}_{s\leq t}\right)$. This follows from the fact that $\tilde{\mathbf{q}}$ depends continuously on η and that the Ornstein–Uhlenbeck process can be expressed as continuous function of the driving noise. This reformulation of (9) was first introduced in [11], where the associated random dynamical system is studied and the existence of an absorbing set is demonstrated.

On the Hilbert space $\mathbf{H}^{-1}$ we consider the Borel σ–algebra $\mathcal{B}(\mathbf{H}^{-1})$. We denote by $B_b(\mathbf{H}^{-1})$ the space of real bounded Borel measurable functions on

$\mathbf{H}^{-1}$, by $C_b(\mathbf{H}^{-1})$ the continuous bounded functions, and by $\mathcal{M}_1(\mathbf{H}^{-1})$ the space of probability measures on $\mathbf{H}^{-1}$. Given the uniqueness of the solutions by Theorem 2.4, it can be shown that $\{\mathbf{q}(t, \mathbf{q}_0),\, t \geq 0\}$ is a Markov process (see e.g. Theorem 9.14 in [12]), where $\mathbf{q}(t, \mathbf{q}_0)$ is the solution seen as a random variables, with the randomness introduced by the driving Brownian motion. For any $\mathbf{q}_0 \in \mathbf{H}^{-1}$ and $A \in \mathcal{B}(\mathbf{H}^{-1})$, and $t > 0$, define the associated Markov transition probabilities as

$$P_t(\mathbf{q}_0, A) = \mathbb{P}\left(\mathbf{q}(t, \mathbf{q}_0) \in A\right) = \mathrm{Law}(\mathbf{q}(t, \mathbf{q}_0))(A).$$

The corresponding Markov semigroup is defined as

$$\mathcal{P}_t\varphi(\mathbf{q}_0) = \int \varphi(x) P_t(\mathbf{q}_0, dx) = \mathbb{E}\,\varphi(\mathbf{q}(t, \cdot; \mathbf{q}_0))$$

for any $\varphi \in B_b(\mathbf{H}^{-1})$, and $\mathcal{P}_t^*$ is its dual acting on $\mathcal{M}_1$ i.e.

$$\mathcal{P}_t^*\mu(A) = \int \mathcal{P}_t \mathbb{1}_A(x)\, \mu(dx) = \int P_t(x, A)\, \mu(dx).$$

Given the continuous dependence of $\mathbf{q}$ on the initial condition it follows that the semigroup $\mathcal{P}_t$ is *Feller* namely that for any $\varphi \in C_b(\mathbf{H}^{-1})$ and any $t \geq 0$ one has $\mathcal{P}_t\varphi \in C_b(\mathbf{H}^{-1})$.

3 Methodology

We discuss a generalised Harris theorem for Markov processes in a Hilbert space. The presentation closely follows [24, Section 4] and, building on a framework developed in [6], we provide conditions particularly suitable for application to dissipative SPDEs.

3.1 Basic definitions and notation

Let $\mathcal{H}$ be a Hilbert space with scalar product $(\cdot,\, \cdot)$ and associated norm $|\cdot|$ and Borel σ–algebra $\mathcal{B}(\mathcal{H})$. We denote by $B_b(\mathcal{H})$ the space of all real bounded Borel functions and by $\mathcal{M}_1(\mathcal{H})$ the space of probability measures on $\mathcal{H}$. Consider two probability measures $\mu_1, \mu_2 \in \mathcal{M}_1(\mathcal{H})$. A probability measure Γ on $\mathcal{H} \times \mathcal{H}$ is called *coupling* of μ_1, μ_2 if its marginals agree with μ_1, μ_2; more specifically, if we let $\pi_1(x, y) = x$, $\pi_2(x, y) = y$ be the projections of $\mathcal{H} \times \mathcal{H}$ onto its two components, then

$$\pi_1^*\Gamma = \mu_1, \quad \text{and} \quad \pi_2^*\Gamma = \mu_2.$$

We denote the set of all such couplings as $\mathcal{C}(\mu_1, \mu_2)$. Equivalently we call a pair of random variables (ξ_1, ξ_2) a coupling of μ_1, μ_2 if $\mathrm{Law}\,\xi_1 = \mu_1$ and $\mathrm{Law}\,\xi_2 = \mu_2$.

A function $d : \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ is a *semimetric* (sometimes also referred to as a *distance–like function*) when it is symmetric, lower semi-continuous and such that $d(x, y) = 0 \Leftrightarrow x = y$. When the symmetry fails, we refer to d as a

premetric. A semimetric d on $\mathcal{H}$ can be lifted to a semimetric on the level of probabilities called *Wasserstein semimetric* W_d: given two probability measures $\mu_1, \mu_2 \in \mathcal{M}_1(\mathcal{H})$ set

$$W_d(\mu_1, \mu_2) := \inf_{\Gamma \in \mathcal{C}(\mu_1, \mu_2)} \int d(x, y)\, \Gamma(dx, dy).$$

The classic coupling lemma (see e.g. [33, Theorem 4.1]) ensures that the infimum in this definition is always reached by a coupling, given the lower semi-continuity of d. Further, note that, if d satisfies the triangular inequality, namely it is a metric, then W_d is a metric on $\mathcal{M}_1$.

An important example is the *total variation distance* $d_{TV}(\mu_1, \mu_2)$ between two probability measures $\mu_1, \mu_2 \in \mathcal{M}_1$, defined as the Wasserstein metric associated with the discrete metric $d(x, y) = \mathbb{1}(x \neq y)$. There is an important connection between Wasserstein semimetrics and weak norms associated with Lipschitz functions. Given a semimetric d, define the associated Lipschitz seminorm as

$$\|\varphi\|_d := \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x, y)}. \tag{23}$$

It is immediate that $\|.\|_d$ is positive homogenous and satisfies the triangle inequality. However, $\|\varphi\|_d = 0$ only implies that φ is a constant function. For any two probability measures $\mu_1, \mu_2 \in \mathcal{M}_1$ it is easy to see that

$$\sup_{\|\varphi\|_d \leq 1} |\langle \varphi, \mu_1 \rangle - \langle \varphi, \mu_2 \rangle| \leq W_d(\mu_1, \mu_2). \tag{24}$$

When d is a metric however, the Kantorovich-Rubinstein formula (see e.g. [16]) states that (24) is an equality.

3.2 A generalisation of Harris' theorem

In a finite dimensional context Harris' theorem (see e.g. [24, Theorem 1.5], [23], [28]) provides conditions under which the transition probabilities of a Markov process converge to an invariant measure in the total variation distance. In fact it ensures convergence if there exists a so-called *small set* which is visited infinitely often by the process, and the speed of convergence is related to how fast the process returns to such sets. A set $A \subset \mathcal{H}$ is *small* if there exists a time $t > 0$ and a constant $\varepsilon > 0$ such that

$$d_{TV}\left(P_t(x, \cdot), P_t(y, \cdot)\right) \leq 1 - \varepsilon \quad \text{for all } x, y \in A.$$

As explained for example in [23], typical candidates for small sets are the level set of so called Lyapunov functions. A measurable function $V : \mathcal{H} \to \mathbb{R}_+$ is called *Lyapunov function* for $\mathcal{P}_t$ if there exist positive constants C_V, γ, K_V such that

$$\mathcal{P}_t V(x) \leq C_V e^{-\gamma t} V(x) + K_V \quad \text{for all } x \in \mathcal{H},\, t \geq 0.$$

However, a major difficulty with applying Harris' theorem in an infinite dimensional context is that the transition probabilities $P_t(x,\cdot)$ and $P_t(y,\cdot)$ might be singular for different initial conditions $x \neq y$. In that case one has that $d_{TV}\left(P_t(x,\cdot), P_t(y,\cdot)\right) = 1$ and there are no small sets. In [24] in order to retrieve a version of Harris' theorem in infinite dimensions, a weaker notion of small set is introduced, one which uses a Wasserstein semimetric, rather than the total variation distance:

Definition 3.1 ([24, Definition 4.4])**.** Let $\mathcal{P}$ be a Markov operator on $\mathcal{H}$ with associated transition function $P(\cdot,\cdot)$ and let $d : \mathcal{H} \times \mathcal{H} \to [0,1]$ be a semimetric on $\mathcal{H}$. A set $K \subset \mathcal{H}$ is called *d-small* for $\mathcal{P}$ if there exists $\varepsilon > 0$ such that

$$W_d(P(x,\cdot), P(y,\cdot)) \leq 1 - \varepsilon$$

for all $x, y \in K$.

Note that the range of the semimetric is restricted to the unit interval. This does not, however, impose a restriction, since if d is a semimetric, then so is $d \wedge 1$, and furthermore $W_{d\wedge 1} \leq W_d$. The reason why the range of d is restricted in this way is that large values of x and y will be dealt with separately using the Lyapunov function, as we will see.

There is a last ingredient necessary to give a general form of Harris' theorem for Wasserstein semimetrics: the semimetric has to be contracting for the semigroup $\{\mathcal{P}_t,\, t \geq 0\}$. In the original statement of Harris' theorem this condition is not included as it is automatically verified by the total variation distance.

Definition 3.2 ([24, Definition 4.6])**.** Let $\mathcal{P}$ be a Markov operator on $\mathcal{H}$ with associated transition function $P(\cdot,\cdot)$. Then a semimetric $d : \mathcal{H} \times \mathcal{H} \to [0,1]$ is called *contracting* for $\mathcal{P}$ if there exists $\alpha < 1$ such that for every pair $x, y \in \mathcal{H}$ with $d(x,y) < 1$

$$W_d(P(x,\cdot), P(y,\cdot)) \leq \alpha\, d(x,y).$$

Going through the proof of Theorem 4.8 in [24], it is clear that first and foremost the following intermediate result is shown:

Theorem 3.3 (Generalised Harris' theorem, [24, Theorem 4.8])**.** *Let $\mathcal{P}_t$, $t \geq 0$, be a Markov semigroup over $\mathcal{H}$ admitting a continuous Lyapunov function V. Suppose that there exists $T > 0$ and a semimetric $d : \mathcal{H} \times \mathcal{H} \to [0,1]$ which is contracting for $\mathcal{P}_T$, and such that the level set $\{x \in \mathcal{H} : V(x) \leq 4K_V\}$ is d-small for $\mathcal{P}_T$. Then the following holds:*

1. *Defining $\tilde{d}(x,y)^2 = d(x,y)(1+V(x)+V(y))$, there exists $t_* > 0$ and $\rho < 1$ such that*
$$W_{\tilde{d}}\left(P_{t_*}(x,\cdot), P_{t_*}(y,\cdot)\right) \leq \rho\, \tilde{d}(x,y). \tag{25}$$

2. *The semigroup $\{\mathcal{P}_t,\, t \geq 0\}$ has at most one invariant measure μ_*.*

3. *If there exists a complete metric d_0 on $\mathcal{H}$ such that $\mathcal{P}_t$ is Feller in $(\mathcal{H}, d_0)$ and $d_0 \leq \sqrt{d}$, then the semigroup $\{\mathcal{P}_t,\, t \geq 0\}$ has an invariant measure.*

From this result two important corollaries will follow, namely the exponential convergence of transition probabilities to the invariant measure, Corollary 3.4, and a spectral gap result, Corollary 3.5.

Corollary 3.4. *Suppose the assumptions of Theorem 3.3 hold uniformly for T belonging to an open interval of $\mathbb{R}$. Then Theorem 3.3 implies that there exist $\gamma > 0$ and a $t_0 > 0$ such that given $\mu, \nu \in \mathcal{M}_1$*

$$W_{\tilde{d}}(\mathcal{P}_t\mu, \mathcal{P}_t\nu) \leq e^{-\gamma t} W_{\tilde{d}}(\mu, \nu) \quad \textit{for all } x \in \mathcal{H},\, t \geq t_0.$$

Furthermore, if μ_* is the invariant measure for $\mathcal{P}_t$, there exists $C > 0$ such that

$$W_{\tilde{d}}(P_t(x, \cdot), \mu_*) \leq C(1 + V(x))e^{-\gamma t} \quad \textit{for all } x \in \mathcal{H},\, t \geq t_0.$$

Recall that μ_* being invariant for $\mathcal{P}_t$ means that μ_* is an eigenvector of $\mathcal{P}_t^*$ with eigenvalue 1. Then the next corollary implies that the remaining spectrum is contained in the disk of radius $e^{-\gamma t}$ around the origin, or we may say $\mathcal{P}_t^*$ exhibits a spectral gap.

Corollary 3.5 (Spectral gap)**.** *Suppose all conditions in Theorem 3.3 are met and let $\|\cdot\|_{\tilde{d}}$ be the Lipschitz seminorm* (23) *associated to the semimetric $\tilde{d}$. Then there exists $\rho < 1$ and $t_* > 0$ such that*

$$\|\mathcal{P}_{t_*}\varphi - \langle \varphi, \mu_* \rangle\|_{\tilde{d}} \leq \rho \|\varphi - \langle \varphi, \mu_* \rangle\|_{\tilde{d}}$$

for all $\varphi : \mathcal{H} \to \mathbb{R}$ such that $\|\varphi\|_{\tilde{d}} < \infty$.

Given the formulation of $\tilde{d}$, for observables such that $\|\varphi\|_{\tilde{d}} < \infty$ we have

$$|\varphi(x) - \varphi(y)| \leq \|\varphi\|_{\tilde{d}} \sqrt{d(x, y)(1 + V(x) + V(y))} \quad \text{for all } x \neq y.$$

This implies that the spectral gap holds for functions that are locally Hölder with exponent $\frac{1}{2}$ with respect to the semimetric d. Yet the precise regularity of those observables with respect to the original norm $|\cdot|$ on $\mathcal{H}$ depends on the formulation of the contracting semimetric d. Recently [6] provided a set of conditions to construct a semimetric satisfying the hypothesis in Theorem 3.3. Given this choice of the semimetric, it turns out that the spectral gap holds for functions that are locally Hölder with respect to the original norm $|\cdot|$ and some exponent $\alpha < \frac{1}{2}$. In the next subsection we propose a modification of those conditions in [6] to facilitate their application to models like the stochastic 2LQG.

3.3 Framework for SPDEs

The following setup includes the dissipative SPDE's we are interested in but potentially other interesting dynamic models in infinite dimensions. Let $(\mathcal{H}, |\cdot|)$ and $(\mathcal{V}, \|\cdot\|)$ be Hilbert spaces with $\mathcal{V} \subset\subset \mathcal{H}$ (i.e. $\mathcal{V}$ is compactly contained in $\mathcal{H}$ or the unit sphere of $\mathcal{V}$ is relatively compact in $\mathcal{H}$). Further, $\|v\| \geq |v|$, and

$\mathcal{V}$ is assumed dense in $\mathcal{H}$. This implies that $\mathcal{H} = \mathcal{H}' \subset\subset \mathcal{V}'$, that $|v| \geq \|v\|'$, and that $\mathcal{H}$ is dense in $\mathcal{V}'$. Consider the stochastic equation

$$dX = (AX + F(X))\, dt + dW_X, \quad X(0) = x \tag{26}$$

where $A : \mathcal{V} \to \mathcal{V}'$ is a nonnegative linear operator, W_X is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathcal{H}$ and trace class covariance operator $Q : \mathcal{H} \to \mathcal{H}$, and $F : \mathcal{V} \to \mathcal{V}'$ is a nonlinear continuous function such that Equation (26) holds in $\mathcal{V}'$, and that there exists a unique solution for any initial condition $X(0) = x \in \mathcal{H}$. As for the two–layer quasi–geostrophic equations, we assume the solution to be in $C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$ for all $T > 0$ and a.a. $\omega \in \Omega$, and to be continuous with respect to the initial condition (as a function into $C([0,T];\mathcal{H})$). We define the Markov semigroup $\mathcal{P}_t$ on $\mathcal{B}_b(\mathcal{H})$ by $\mathcal{P}_t\varphi(x) = \mathbb{E}\,\varphi(X(t;x))$ and denote the associated transition probabilities as $P_t(x,\cdot)$. Given the regularity of the solutions the associated semigroup is Feller.

To ensure the generalised Harris' theorem applies we need to find a semimetric d which is contracting with respect to $\mathcal{P}_t$ for some t and show that there is a Lyapunov function V with a d-small sublevel set. By definition then we have to appropriately bound the distance $W_d(P_t(x,\cdot), P_t(y,\cdot))$, for x, y such that $d(x,y) < 1$ in order to get the contraction, and for x, y in a sublevel set of V in order to get the d–smallness.

Let Y satisfy the stochastic equation

$$dY = (AY + F(Y))\, dt + dW_Y, \quad Y(0) = y \tag{27}$$

with A, F as in Equation (26), while W_Y and W_X have the same law but are not necessarily identical (note also the different initial condition $Y(0) = y \neq x$). As a consequence $\operatorname{Law} Y(t) = P_t(y,\cdot)$, so any coupling between W_Y and W_X furnishes a coupling between $X(t)$ and $Y(t)$ (or equivalently between $P_t(x,\cdot)$ and $P_t(y,\cdot)$) implying that $\mathbb{E}(d(X(t),Y(t)))$ is an upper bound for $W_d(P_t(x,\cdot), P_t(y,\cdot))$. Taking the discrete metric $d(x,y) = \mathbb{1}(x \neq y)$ for instance, we find $1 - \mathbb{P}(X(t) = Y(t))$ as an upper bound for the total variation metric. The event $X(t) = Y(t)$ though implies, roughly speaking, that the process X has forgotten its initial condition. Finding a coupling however that guarantees this with nontrivial probability is difficult in our infinite–dimensional setting. While memory of the initial condition will eventually get wiped out in those degrees of freedom directly affected by the noise, the required time might get very long if there are infinitely many such degrees of freedom. Clearly, this memory might also survive in degrees of freedom where there is no noise at all.

The idea of *generalised coupling* is to consider an intermediate equation

$$d\tilde{Y} = \Big(A\tilde{Y} + F(\tilde{Y}) + G(X,\tilde{Y})\Big)\, dt + dW_X, \quad \tilde{Y}(0) = y \tag{28}$$

where A, F and even W_X are as in Equation (26), but with initial condition y as in Equation (27). Further, G is a control function determined in such a way that on the one hand for a t large enough we can estimate $\mathbb{E}(d(X(t),\tilde{Y}(t)))$ even if $x \neq y$ (using an appropriate semimetric d). On the other hand G has to be

chosen so that, under appropriate conditions, the law of $\tilde{Y}$ will be absolutely continuous with respect to the law of Y and so that the total–variation metric between the laws of $\tilde{Y}$ and Y can be controlled (as in Lemma 3.9). We will achieve this by choosing G such that the law of the process

$$\tilde{W}_X(t) = \int_0^t G(X, \tilde{Y})\, ds + W_X(t) \tag{29}$$

is absolutely continuous with respect to the law of W_Y.

Let us now collect the assumptions we will be using and which are particularly suitable for applications to dissipative SPDEs.

Assumption A. Let $\mathcal{H}_n$ be an n-dimensional subspace of $\mathcal{H}$ and denote by Π_n the orthogonal projection on $\mathcal{H}_n$. The covariance operator Q commutes with Π_n, and furthermore $Q_n := \Pi_n Q$ is invertible on $\mathcal{H}_n$. Given a solution $t \to X(t)$ of (26), there exists a finite dimensional measurable control $G : \mathcal{H} \times \mathcal{H} \to \mathcal{H}_n$ such that the controlled equation (28) has a unique solution $\tilde{Y}$ in the sense described at the beginning of this subsection. In addition, we require the following:

A1 There exist $\kappa_0 > 0$ and $\kappa_1 \geq 0$ such that for all $t \geq 0$

$$|X(t) - \tilde{Y}(t)|^2 \leq |x - y|^2 \exp\left(-\kappa_0 t + \kappa_1 \int_0^t \|X(s)\|^2\, ds\right).$$

A2 There exists $\kappa_2 > 0$, $\kappa_3 \geq 0$ and a random variable Ξ_γ such that

$$|X(t)|^2 + \kappa_2 \int_0^t \|X(s)\|^2\, ds \leq |x|^2 + \kappa_3 t + \Xi_\gamma \quad t \geq 0 \tag{30}$$

with $\kappa_0 > \kappa_1\kappa_3/\kappa_2$ and

$$\mathbb{P}(\Xi_\gamma \geq R) \leq e^{-2\gamma R}, \quad R \geq 0. \tag{31}$$

A3 There exists a positive constant $c > 0$ such that for each $t \geq 0$ and $s \in [0, t]$

$$|G(X(s), \tilde{Y}(s))|^2 \leq c|X(s) - \tilde{Y}(s)|^2.$$

A4 There exists a measurable function $V : \mathcal{H} \to \mathbb{R}_+$ such that for some $\gamma_1 > 0$, $K > 0$

$$\mathbb{E}V(X(t)) \leq \mathbb{E}V(X(s)) + \int_s^t (-\gamma_1 \mathbb{E}V(X(\tau)) + K)\ d\tau, \quad t \geq s \geq 0. \tag{32}$$

Furthermore for any $M > 0$ the function $x \mapsto |x|^2$ is bounded on the level sets $\{V \leq M\}$.

Given κ_1, κ_2 from **A1** and **A2** respectively define the premetric θ_α depending on a positive parameter α as

$$\theta_\alpha(x,y) := |x-y|^{2\alpha} e^{\alpha \upsilon |x|^2} \quad \text{with } \upsilon := \frac{\kappa_1}{\kappa_2} \tag{33}$$

and, given $N \in \mathbb{N}$, define a semimetric d_N as

$$d_N(x,y) := N\theta_\alpha(x,y) \wedge N\theta_\alpha(y,x) \wedge 1. \tag{34}$$

We will see in Theorem 3.7 that, thanks to Assumption A, there exists α_0 such that for all $\alpha \in (0, \alpha_0)$ the conditions of Theorem 3.3 are satisfied with respect to the associated semimetric d_N for large enough N. First though we use Assumption A to demonstrate an important bound on the control term G in Equation (28). This will be exploited in Lemma 3.9 which we take from [6].

Lemma 3.6. *Let the processes X and $\tilde{Y}$ be the solutions of Equation* (26) *resp. of the controlled equation* (28)*. If Assumptions **A1**-**A3** hold, then for all $t \geq 0$*

$$\int_0^t |Q_n^{-1/2} G(X(s), \tilde{Y}(s))|^2 \, ds \leq \tfrac{c\|Q_n^{-1/2}\|^2}{\chi} |x-y|^2 e^{\frac{\kappa_1}{\kappa_2}\left(|x|^2 + \Xi_\gamma\right)} \left(1 - e^{-\chi t}\right) \tag{35}$$

for some $\chi > 0$.

Proof. By **A3** and **A1** we have that for any $t \geq 0$ and $s \in [0,t]$

$$|G(X(s), \tilde{Y}(s))|^2 \leq c|x-y|^2 \exp\Big(-\kappa_0 s + \kappa_1 \int_0^s \|X(\tau)\|^2 \, d\tau\Big),$$

and by **A2**

$$|G(X(s), \tilde{Y}(s))|^2 \leq c|x-y|^2 \exp\Big(-\Big(\kappa_0 - \tfrac{\kappa_1\kappa_3}{\kappa_2}\Big) s + \tfrac{\kappa_1}{\kappa_2}\left(|x|^2 + \Xi_\gamma\right)\Big).$$

It follows that for any $t \geq 0$

$$\begin{aligned}\int_0^t |Q_n^{-1/2} G(X(s), \tilde{Y}(s))|^2 \, ds &\leq \|Q_n^{-1/2}\|^2 \int_0^t |G(X(s), Y(s))|^2 \, ds \\ &\leq \tfrac{c\|Q_n^{-1/2}\|^2}{\chi} |x-y|^2 \exp\Big(\tfrac{\kappa_1}{\kappa_2}\left(|x|^2 + \Xi_\gamma\right)\Big)\left(1 - e^{-\chi t}\right)\end{aligned}$$

where $\chi = \kappa_0 - \frac{\kappa_1\kappa_3}{\kappa_2} > 0$. □

Given n and Q_n as in Assumption A, set $W_n := \Pi_n W_X$, which is a Wiener process with covariance matrix Q_n and independent of $(1-\Pi_n)W_X$. A consequence of Lemma 3.6 is that the process

$$\tilde{W}_n(t) := W_n(t) + \int_0^t G(X(s), Y(s)) \, ds$$

is absolutely continuous with respect to W_n, by Girsanov's theorem. Then $\tilde{W}_X$ in Equation (29), i.e. $\tilde{W}_X = \tilde{W}_n + (1-\Pi_n)W_X$ is a Q-Wiener process absolutely continuous with respect to W_X, hence also with respect to W_Y as $\operatorname{Law} W_X = \operatorname{Law} W_Y$.

We are ready to state the main theorem of this section.

Theorem 3.7. *Consider $X(t)$ solution of (26) with associated Markov semigroup $\mathcal{P}_t$ which is Feller in $(\mathcal{H}, |\cdot|)$ and Assumption A holds. Given κ_1, κ_2 and γ as in **A1** and **A2**, set*

$$\upsilon = \frac{\kappa_1}{\kappa_2} \quad \text{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma}. \tag{36}$$

Then there exists a unique invariant measure μ_*, and there exists $t > 0$ and $\rho \in (0,1)$ such that

$$W_{\tilde{d}}(P_t(x,\cdot), P_t(y,\cdot)) \leq \rho\, \tilde{d}(x,y). \tag{37}$$

where $\tilde{d}(x,y)^2 = d_N(x,y)\,(1 + V(x) + V(y))$ and d_N is defined as in (34) for υ and α_0 as above.

In order to prove this result we introduce first the following three lemmas. The first one is the so-called comparison theorem, a classic generalisation of the integral Gronwall lemma:

Lemma 3.8 (Comparison Theorem)**.** *Let f be a continuous function, for which*

$$f(s) - f(r) \leq -\gamma \int_r^s f(\tau)\, d\tau + K(s-r) \quad \text{for all } 0 < r < s \tag{38}$$

with $\gamma, K > 0$ then

$$f(t) \leq f(0)e^{-\gamma t} + K/\gamma \quad \text{for all } t \geq 0. \tag{39}$$

The next lemma gives upper bounds for the total variation distance between a finite dimensional Wiener process and the same process with added drift.

Lemma 3.9 ([6, Theorem A.5])**.** *Let B be a n-dimensional Wiener process with covariance operator σ, $(h(t))_{t\geq 0}$ a progressively measurable n-dimensional process and define*

$$\tilde{B}(t) = \int_0^t h(s)\, ds + B(t) \quad t \geq 0.$$

Fix $t > 0$, then, if for some $\delta \in (0,1)$

$$M_\delta := \mathbb{E}\left(\int_0^t |\sigma^{-1/2} h(s)|^2\, ds\right)^\delta < \infty \tag{40}$$

the following bounds hold:

$$d_{TV}(\mathrm{Law}(B(s))_{s\leq t}, \mathrm{Law}(\tilde{B}(s))_{s\leq t}) \leq 2^{(1-\delta)/(1+\delta)} M_\delta^{1/(1+\delta)}; \tag{41}$$

$$d_{TV}(\mathrm{Law}(B(s))_{s\leq t}, \mathrm{Law}(\tilde{B}(s))_{s\leq t}) \leq 1 - \frac{1}{6}\min\left(\frac{1}{8}, \exp\left(-(2^{2-\delta}M_\delta)^{1/\delta}\right)\right). \tag{42}$$

The third lemma provides an estimate for the distance of two transition probabilities with different initial condition which will be the first block in building the proof of Theorem 3.7. To state the lemma, we set $n \in \mathbb{N}$ as in Assumption A and recall the definition

$$W_n = \Pi_n W_X \qquad \text{and} \qquad \tilde{W}_n = \Pi_n \tilde{W}_X. \tag{43}$$

Lemma 3.10. *Let $X(t)$ and $Y(t)$ be the solutions of (26) and (27) respectively and $P_t(x,\cdot), P_t(y,\cdot)$ their respective laws. If Assumption A holds, then there exists positive constants C_Ξ and α_0 such that for all $\alpha \in (0,\alpha_0)$ and $N \in \mathbb{N}$ we have*

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \leq d_{TV}(\mathrm{Law}(W_n(t))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(t))_{s\leq t}) + NC_\Xi \theta_\alpha(x,y) e^{-\chi\alpha t}, \tag{44}$$

for all $x, y \in \mathcal{H}$.

Proof. As shown in [6, Theorem 2.4], by means of the classic coupling lemma and the gluing lemma (see e.g. [33]) it can be shown that

$$W_{d_N}(P_t(x,.), P_t(y,.)) \leq \underbrace{d_{TV}(P_t(y,\cdot), \mathrm{Law}\,\tilde{Y}(t))}_{(I)} + \underbrace{\mathbb{E}[d_N(X(t), \tilde{Y}(t))]}_{(II)}. \tag{45}$$

We look at the two terms separately starting from (I). From the theory of stochastic differential equations the solutions of (27) and (28) can be seen as image via a measurable function Φ^y of their driving noise, W_Y and $\tilde{W}_X$, respectively i.e. $Y(t) = \Phi^y((W_Y(s))_{s\leq t})$ and $\tilde{Y} = \Phi^y((\tilde{W}_X(s))_{s\leq t})$. Then, using a classical property of the total variation norm, we get

$$d_{TV}(\mathrm{Law}\,Y(t), \mathrm{Law}\,\tilde{Y}(t)) \leq d_{TV}(\mathrm{Law}(W_Y(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_X(s))_{s\leq t}).$$

By the coupling lemma there exists a coupling $(\xi, \tilde{\xi})$ of the laws of the finite dimensional Wiener processes $(W_n(s))_{s\leq t}$ and $(\tilde{W}_n(s))_{s\leq t}$ such that

$$d_{TV}(\mathrm{Law}(W_n(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(s))_{s\leq t}) = \mathbb{P}(\xi \neq \tilde{\xi}). \tag{46}$$

Therefore $B := \xi + (1-\Pi_n)W_X$ and $\tilde{B} = \tilde{\xi} + (1-\Pi_n)W_X$ provide a coupling for $(W_Y, \tilde{W}_X)$. Hence

$$d_{TV}(\mathrm{Law}(W_Y(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_X(s))_{s\leq t}) \leq \mathbb{P}(B \neq \tilde{B}). \tag{47}$$

But since $B = \tilde{B} \Leftrightarrow \xi = \tilde{\xi}$ we can use Equation (46) in Equation (47) and obtain

$$d_{TV}(\mathrm{Law}(W_Y(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_X(s))_{s\leq t}) \leq d_{TV}(\mathrm{Law}(W_n(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(s))_{s\leq t})$$

giving the first part of (44).

Next we look at (II) in (45). By definition of d_N (34) we have

$$\mathbb{E}\, d_N(X(t), \tilde{Y}(t)) \leq N\mathbb{E}\, \theta_\alpha(X(t), \tilde{Y}(y)). \tag{48}$$

Combining Assumption **A1** and **A2** we have that

$$|X(t) - \tilde{Y}(t)|^2 \leq |x-y|^2 \exp\big(-\chi t + \upsilon(|x|^2 - |X(t)|^2 + \Xi_\gamma)\big) \tag{49}$$

namely, for all $\alpha > 0$

$$|X(t) - \tilde{Y}(t)|^{2\alpha} e^{\alpha\upsilon|X(t)|^2} \leq |x-y|^{2\alpha} e^{\alpha\upsilon|x|^2} \exp(-\alpha\chi t + \alpha\upsilon\Xi_\gamma).$$

Therefore by the definition of the premetric θ_α (33)

$$\mathbb{E}\,\theta_\alpha(X(t), \tilde{Y}(t)) \leq C_\Xi \theta_\alpha(x,y) \exp(-\chi\alpha t) \tag{50}$$

where by (31)

$$C_\Xi := \mathbb{E}\,\exp(\upsilon\alpha\Xi_\gamma) < \infty \quad \text{if } \upsilon\alpha < 2\gamma. \tag{51}$$

Putting together these results in (45) we have that (44) holds for all $\alpha \in (0, \alpha_0)$ setting $\alpha_0 := 2\gamma/\upsilon$.

□

We are now ready to show Theorem 3.7:

Proof of Theorem 3.7. First we show that V is a Lyapunov function as in (3.2). By **A4** we know that there exists γ_1, K strictly positive such that

$$\mathcal{P}_t V(x) - \mathcal{P}_s V(x) \leq \int_s^t \left(-\gamma_1 \mathcal{P}_\tau V(x) + K\right) d\tau, \quad t \geq s \geq 0. \tag{52}$$

Therefore the hypothesis of Lemma 3.8 are satisfied by the function $f(t) := \mathcal{P}_t V(x)$, which has non-negative values and is continuous in time, thus we have the desired result

$$\mathcal{P}_t V(x) \leq e^{-\gamma_1 t} V(x) + K_V \tag{53}$$

with $K_V = K/\gamma_1$.

Our result will follow from the Generalised Harris' Theorem 3.3 if we are able to show that for all $\alpha \in (0, \alpha_0)$, there exists $N_* \in \mathbb{N}$ and $t_* \in \mathbb{R}_+$ such that for all $t > t_*$ and $N \in \mathbb{N}$ such that $N > N_*$:

(i) the semimetric d_N (34) is contracting for $\mathcal{P}_t$;

(ii) the level set $\{x \in \mathcal{H} : V(x) \leq 4K_V\}$ is d_N-small.

(iii) There exists a complete metric d_0 such that $d_0 \leq \sqrt{d_N}$, and further $\mathcal{P}_t$ is Feller on $(\mathcal{H}, d_0)$.

We start by showing item (iii). The Hilbert space $\mathcal{H}$ is endowed with the norm $|\cdot|$ associated to its scalar product. Define $d_0(x,y) = |x-y| \wedge 1$. This is a complete metric that defines the same topology as the norm $|\cdot|$, and as $\mathcal{P}_t$ is Feller in $(\mathcal{H}, |\cdot|)$, it is in particular Feller in $(\mathcal{H}, d_0)$. Furthermore, we have

$$\sqrt{d_N(x,y)} \geq N^{1/2}|x-y|^\alpha \wedge 1 \geq |x-y| \wedge 1 = d_0(x,y)$$

as $\alpha \in (0, \alpha_0)$ with $\alpha_0 \leq 1/2$. The remainder of the proof is dedicated to show the properties (i) and (ii).

(i) The semigroup $\mathcal{P}_t$ is d_N–contracting. We have to show that there exists $\rho < 1$ such that

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \leq \rho\, d_N(x,y)$$

for all x, y such that $d_N(x,y) < 1$. Thanks to Lemma 3.10 we have that for all $\alpha < 2\gamma/\upsilon$

$$W_{d_N}(P_t(x,.), P_t(y,.)) \leq d_{TV}(\mathrm{Law}(W_n(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(s))_{s\leq t}) + NC_\Xi \theta_\alpha(x,y) e^{-\chi\alpha t} \quad (54)$$

We want to apply Lemma 3.9 to bound the first term on the right hand side so we have to ensure that (40) holds. In (29) we defined

$$\tilde{W}_X(t) = W_X(t) + \int_0^t G(X, \tilde{Y})\, ds$$

and in particular, since G is n-dimensional, we have

$$\Pi_n \tilde{W}_X(t) = \Pi_n W_X(t) + \int_0^t G(X, \tilde{Y})\, ds.$$

Then, using Q_n covariance operator of the Wiener process $W_n = \Pi_n W_X$ as in (22), in order to apply Lemma 3.9 we have to show that

$$M_\delta = \mathbb{E}\left(\int_0^t |Q_n^{-1/2} G(X, \tilde{Y})|^2\, ds\right)^\delta \quad (55)$$

is finite. By Lemma 3.6 one has

$$\int_0^t |Q_n^{-1/2} G(X, \tilde{Y})|^2\, ds \leq \tfrac{c\|Q_n^{-1/2}\|^2}{\chi} |x-y|^2 \exp\big(\upsilon\left(|x|^2 + \Xi_\gamma\right)\big)\left(1 - e^{-\chi t}\right)$$

where

$$\upsilon := \frac{\kappa_1}{\kappa_2} \quad \text{and} \quad \chi = \kappa_0 - \upsilon\kappa_3 > 0.$$

Since Equation (31) holds for Ξ_γ we have that

$$\mathbb{E}\, \exp(\upsilon\delta\Xi_\gamma) < \infty \quad \text{if } \upsilon\delta < 2\gamma$$

and so for all $0 < \delta < (2\gamma/\upsilon) \wedge 1$

$$M_\delta \leq \tilde{C}_\delta |x-y|^{2\delta} \exp\big(\upsilon\delta|x|^2\big) \quad \text{with } \tilde{C}_\delta = \left(\tfrac{c\|Q_n^{-1/2}\|^2}{\chi}\right)^\delta \mathbb{E}\, e^{\upsilon\delta\Xi_\gamma}. \quad (56)$$

Therefore, since condition (40) holds, the bound (41) in Lemma 3.9 and (56) give

$$\begin{aligned} d_{TV}(\mathrm{Law}(W_n(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(s))_{s\leq t}) &\leq 2^{(1-\delta)/(1+\delta)} M_\delta^{\frac{1}{1+\delta}} \\ &\leq C_{\frac{\delta}{1+\delta}} \left(|x-y|^2 \exp\big(\upsilon|x|^2\big)\right)^{\frac{\delta}{1+\delta}} \end{aligned}$$

where $C_{\frac{\delta}{1+\delta}} = 2^{(1-\delta)/(1+\delta)}\tilde{C}_\delta$. Since $\delta \in (0, (2\gamma/\upsilon)\wedge 1)$, the exponent $\delta/(1+\delta) =: \alpha$ is in the interval

$$0 < \alpha < \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma} < \frac{2\gamma}{\upsilon}.$$

Given the definition of the premetric θ_α (33) we have then shown that for all $\alpha \in (0, \alpha_0)$ there exists $C_\alpha > 0$ such that

$$d_{TV}(\mathrm{Law}(W_n(s))_{s\le t}, \mathrm{Law}(\tilde{W}_n(s))_{s\le t}) \le C_\alpha \theta_\alpha(x,y) \quad \text{for all } x, y \in \mathcal{H},\ x \neq y.$$

Going back to (54), we have proved that

$$\begin{aligned} W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) &\le C_\alpha \theta_\alpha(x,y) + NC_\Xi e^{-\chi\alpha t}\theta_\alpha(x,y) \\ &= N\theta_\alpha(x,y)\left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t}\right) \end{aligned}$$

and, inverting the roles of x and y, we get in the same way

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \le N\theta_\alpha(y,x)\left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t}\right).$$

Therefore

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \le (N\theta_\alpha(x,y) \wedge N\theta_\alpha(y,x))\left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t}\right) \quad (57)$$

and if $d_N(x,y) < 1$ we have that

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \le d_N(x,y)\left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t}\right).$$

Then for all $N \in \mathbb{N}$ and $t > 0$ such that

$$\rho := C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} < 1 \quad (58)$$

we showed that $\mathcal{P}_t$ is d_N-contracting.

(ii) The sublevel set of V is d_N-small. We have now to prove that there exists $\varepsilon > 0$ such that

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \le 1 - \varepsilon \quad (59)$$

for all $x, y \in \{z \in \mathcal{H} : V(z) \le 4K_V\}$. Although a closer examination of the previous step will show that (57) holds in fact for all x, y (not only for x, y such that $d_N(x,y) < 1$), there is no way to obtain (59) directly from (57). Despite $\theta_\alpha(x,y)$ being bounded over the sublevel set $\{V \le 4K_V\}$ (see Eq. 60 below), as soon as $\theta_\alpha(x,y) \wedge \theta_\alpha(y,x) > 0$, we would have to make $C_\alpha + NC_\Xi e^{-\chi\alpha t}$ arbitrarily small to have (59), which is impossible.

However we can use Equation (42) in Lemma 3.9 to arrive at a different estimate for $d_{TV}(\mathrm{Law}(W_n(s))_{s\le t}, \mathrm{Law}(\tilde{W}_n(s))_{s\le t})$ in (54). Using that result, we have that

$$d_{TV}(\mathrm{Law}(W_n(s))_{s\le t}, \mathrm{Law}(\tilde{W}_n(s))_{s\le t}) \le 1 - \frac{1}{6}\min\left(\frac{1}{8}, \exp\left(-(2^{2-\delta}M_\delta)^{1/\delta}\right)\right),$$

where M_δ is as in (55). Since, by Assumption **A4**, $z \mapsto |z|^2$ is bounded over $\{V \leq 4K_V\}$, so is $\theta_\alpha(x,y)$ i.e. there exists $C_K > 0$ such that

$$\theta_\alpha(x,y) = |x-y|^{2\alpha} e^{\alpha \upsilon |x|^2} < C_K \tag{60}$$

for all $x, y \in \{V \leq 4K_V\}$. Thanks to (56) and (60) it follows that

$$M_\delta \leq \tilde{C}_\delta |x-y|^{2\delta} \exp\big(\upsilon\delta|x|^2\big) \leq \tilde{C}_\delta C_K$$

for all $0 < \delta < (2\gamma/\upsilon) \wedge 1$. Setting $\varepsilon_1(\delta)$ to be

$$\varepsilon_1 = \frac{1}{6} \min\left(\frac{1}{8}, \exp\Big(-(2^{2-\delta}\tilde{C}_\delta C_K)^{1/\delta}\Big)\right),$$

we have that

$$d_{TV}(\mathrm{Law}(W_n(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(s))_{s\leq t}) \leq 1 - \varepsilon_1. \tag{61}$$

Then from (54) i.e.

$$W_{\tilde{d}}(P_t(x,\cdot), P_t(y,\cdot)) \leq d_{TV}(\mathrm{Law}(W_n(s))_{s\leq t}, \mathrm{Law}(\tilde{W}_n(s))_{s\leq t}) + NC_\Xi e^{-\alpha\chi t}\theta_\alpha(x,y)$$

by (60) and (61) we have

$$W_{\tilde{d}}(P_t(x,\cdot), P_t(y,\cdot)) \leq 1 - \varepsilon_1 + NC_K C_\Xi e^{-\chi\alpha t}.$$

Then for all $N \in \mathbb{N}$ and $t > 0$ such that

$$\varepsilon_1 - NC_K C_\Xi e^{-\chi\alpha t} =: \varepsilon > 0 \tag{62}$$

we showed that the level set $\{x \in \mathcal{H} : V(x) \leq 4K_V\}$ is d_N-small. The proof is completed by noting that there are N_*, t_* so that conditions (58) and (62) are satisfied simultaneously. □

Thanks to the result just shown one has a precise formulation of the spectral gap property Corollary 3.5, crucial ingredient to develop response theory. In fact by Theorem 3.7 and Theorem 3.3 we know that such a distance–like function is $\tilde{d}(x,y)^2 = d_N(x,y)(1 + V(x) + V(y))$, with d_N as in (34) and V a Lyapunov function of the system. Note that, by definition, the distance-like function d_N is comparable to the α_0–power of the original metric on the space, with α_0 as in Theorem 3.7. Therefore $\mathcal{P}_t$, the semigroup associated to the solution of (26), exhibits a spectral gap on the set of observables which are α_0–Hölder continuous over the level sets of the Lyapunov function V.

4 Spectral Gap for the stochastic 2LQG model

Let $\mathbf{q}(t, \mathbf{q}_0)$ be the solution of the stochastic two–layer quasi–geostrophic model (9) on the Hilbert space $\mathcal{H} = (\mathbf{H}^{-1}, \|\cdot\|_{-1})$ and $\mathcal{P}_t$ the associated semigroup i.e.

$$\mathcal{P}_t\varphi(\mathbf{q}_0) = \mathbb{E}\,\varphi(\mathbf{q}(t,\mathbf{q}_0)) \quad \text{for all } \varphi \in B_b(\mathcal{H}).$$

The semigroup is Feller in the metric introduced by $|||\cdot|||_{-1}$ since the solution can be shown to be continuous in the initial conditions with respect to the metric induced by $|||\cdot|||_{-1}$ (see [7]). Then we want to ensure that this system has exponential ergodicity as described in Section 3, namely there exists a unique invariant measure μ_*, there exists $C > 0$, $\gamma > 0$ such that

$$W_{\tilde{d}}\left(P_t(x,\cdot),\mu_*\right) \leq C\left(1+V(x)\right)e^{-\gamma t} \quad \text{for all } x \in \mathcal{H},\, t \geq 0$$

and $\mathcal{P}_t$ has a spectral gap i.e. there exists $\rho < 1$ such that

$$\|\mathcal{P}_t\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} \leq \rho\|\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}}$$

for all $\varphi : \mathcal{H} \to \mathbb{R}$ with $\|\varphi\|_{\tilde{d}} < \infty$. We consider $\tilde{d}$ as in Theorem 3.3 and d_N as in (34), now in $\mathcal{H} = \left(\mathbf{H}^{-1}, |||\cdot|||_{-1}\right)$, namely

$$\begin{aligned} \tilde{d}(\mathbf{x},\mathbf{y})^2 &= d_N(\mathbf{x},\mathbf{y})(1+V(\mathbf{x})+V(\mathbf{y})) \quad \text{with} \\ d_N(\mathbf{x},\mathbf{y}) &= N\theta_\alpha(\mathbf{x},\mathbf{y}) \wedge N\theta_\alpha(\mathbf{x},\mathbf{y}) \wedge 1 \quad \text{and} \\ \theta_\alpha(\mathbf{x},\mathbf{y}) &= e^{\alpha\upsilon|||\mathbf{x}|||_{-1}^2}|||\mathbf{x}-\mathbf{y}|||_{-1}^{2\alpha}, \quad \alpha \in (0,\alpha_0) \end{aligned} \tag{63}$$

for $\mathbf{x},\mathbf{y} \in \mathbf{H}^{-1}$ and an appropriate choice of the parameters N, υ and α_0 given in the following main theorem of this section. Further, $\{\lambda_n\}_{n\in\mathbb{N}}$ is an increasing sequence of eigenvalues of $-\Delta$.

Theorem 4.1. *Given the stochastic 2LQG model* (9)*, there exists* r_0 *(depending on* ν, Q *and the forcing* f*, see Eq.* (76)*) with the following properties: Suppose that*

(i) $r > r_0$*, and*

(ii) there exists $n \in \mathbb{N}$ *such that* $\Pi_n\mathcal{H} \subset \operatorname{range} Q$*, and*

(iii) $\nu - 2r\lambda_n^{-1} > 0$*.*

Then there exists a unique invariant measure μ_* *as well as* $t > 0$ *and* $\rho < 1$ *such that*

$$W_{\tilde{d}}\left(P_t(\mathbf{q}_0,\cdot),P_t(\tilde{\mathbf{q}}_0,\cdot)\right) \leq \rho\,\tilde{d}(\mathbf{q}_0,\tilde{\mathbf{q}}_0) \tag{64}$$

for all $\mathbf{q}_0,\tilde{\mathbf{q}}_0 \in \mathcal{H}$*. Here* $\tilde{d}$ *is as in* (63) *with Lyapunov function* $V(\mathbf{x}) = |||\mathbf{x}|||_{-1}^2$ *and parameters*

$$\upsilon = \frac{k_B}{\nu - \frac{2\gamma\operatorname{Tr}Q}{\lambda_1^2}} \quad \textit{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma},$$

where $0 < \gamma < \lambda_1^2\nu/2\operatorname{Tr}Q$*,* $k_B := k_0^2/2\nu$ *and* k_0 *as in Lemma 2.3.*

Proof. As the semigroup $\mathcal{P}_t$ is Feller in $\mathcal{H}$ we only have to ensure that Assumption A is satisfied, as then Theorem 3.7 gives the desired result. Consider, beside the original equation (9), the controlled system

$$\begin{aligned} &d\tilde{\mathbf{q}} + \left(B(\tilde{\boldsymbol{\psi}},\tilde{\boldsymbol{\psi}}) + \beta\partial_x\tilde{\boldsymbol{\psi}}\right)dt = \nu\Delta^2\tilde{\boldsymbol{\psi}}\,dt + \begin{pmatrix} f + G(\mathbf{q},\tilde{\mathbf{q}}) \\ -r\Delta\tilde{\psi}_1 \end{pmatrix}dt + d\boldsymbol{W} \\ &\tilde{\mathbf{q}} = (\Delta + M)\tilde{\boldsymbol{\psi}} \end{aligned} \tag{65}$$

with initial condition $\tilde{\mathbf{q}}(0) = \tilde{\mathbf{q}}_0 \neq \mathbf{q}_0$, $\boldsymbol{W} = (W, 0)^t$, and control G to be determined later. Define the difference variables

$$\boldsymbol{\xi} = \mathbf{q} - \tilde{\mathbf{q}} \quad \text{and} \quad \boldsymbol{\phi} = \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}. \tag{66}$$

Then $\boldsymbol{\xi} = (\Delta + M)\boldsymbol{\phi}$ satisfies the equation

$$\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t} + B(\boldsymbol{\phi}, \tilde{\boldsymbol{\psi}}) + B(\boldsymbol{\psi}, \boldsymbol{\phi}) = \nu\Delta^2\boldsymbol{\phi} - \begin{pmatrix} G(\mathbf{q}, \tilde{\mathbf{q}}) \\ r\Delta\phi_2 \end{pmatrix} \tag{67}$$

with initial condition $\boldsymbol{\xi}_0 = \mathbf{q}_0 - \tilde{\mathbf{q}}_0 \neq 0$.

Let $\{\lambda_k\}_{k\in\mathbb{N}}$ be an increasing sequence of eigenvalues of $-\Delta$ with corresponding eigenvectors e_k forming an orthonormal basis for $\mathcal{H}$. Consider the following finite dimensional control

$$G(\mathbf{q}, \tilde{\mathbf{q}}) = a\Pi_n(\Delta\psi_1 - \Delta\tilde{\psi}_1) = a\Pi_n\Delta\phi_1 \tag{68}$$

where Π_n is the projection onto $\mathcal{H}_n = \mathrm{span}\{e_k,\ k = 1, \ldots n\}$ and $a > 0$ is a parameter to be found below. The controlled system is well posed in the sense of Theorem 2.4 as it can be treated effectively as the uncontrolled system. In fact the control can be split into two parts with finite rank, a lower level perturbation of the viscosity term and an additional time dependent forcing which is continuous in time.

Proof of A1. Taking the $\mathbf{L}^2$ scalar product of (67) with $\boldsymbol{\phi}$, we obtain

$$\left(\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}, \boldsymbol{\phi}\right) + (B(\boldsymbol{\psi}, \boldsymbol{\phi}), \boldsymbol{\phi}) = \nu|\Delta\boldsymbol{\phi}|^2 - h_1(a\Pi_n\Delta\phi_1, \phi_1) + rh_2\|\phi_2\|^2,$$

where we have used the fact that $(B(\boldsymbol{\xi}, \boldsymbol{\psi}), \boldsymbol{\xi}) = 0$ of (20). Recall that by (15) we have

$$\left(\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}, \boldsymbol{\phi}\right) = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\xi}, \boldsymbol{\phi}) - \left(\boldsymbol{\xi}, \frac{\mathrm{d}\boldsymbol{\phi}}{\mathrm{d}t}\right) = -\frac{\mathrm{d}|||\boldsymbol{\xi}|||_{-1}^2}{\mathrm{d}t} - \left(-\tilde{A}\boldsymbol{\phi}, \frac{\mathrm{d}\boldsymbol{\phi}}{\mathrm{d}t}\right)$$

and since the operator $\tilde{A}$ is self-adjoint

$$\left(\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}, \boldsymbol{\phi}\right) = -\frac{\mathrm{d}|||\boldsymbol{\xi}|||_{-1}^2}{\mathrm{d}t} - \left(\boldsymbol{\phi}, \frac{\mathrm{d}(-\tilde{A}\boldsymbol{\phi})}{\mathrm{d}t}\right)$$

$$2\left(\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}, \boldsymbol{\phi}\right) = -\frac{\mathrm{d}|||\boldsymbol{\xi}|||_{-1}^2}{\mathrm{d}t}.$$

Then we obtain the following equation

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||_{-1}^2 + \nu|\Delta\boldsymbol{\phi}|^2 + rh_2\|\phi_2\|^2 = (B(\boldsymbol{\psi}, \boldsymbol{\phi}), \boldsymbol{\phi}) + h_1(a\Pi_n\Delta\phi_1, \phi_1).$$

Denoting the orthogonal complement of Π_n as $\Pi_n^\perp$, we can write the control term as $G(\mathbf{q}, \tilde{\mathbf{q}}) = a\Delta\phi_1 - a\Pi_n^\perp\Delta\phi_1$. One can consider the term $a\Pi_n^\perp\Delta\phi_1$ as

an error term because we have to work with finite dimensional controls. Then, given the generalised Poincaré inequalities for any $n \geq 1$ i.e.

$$\|\Pi_n \phi_1\|_{k+1}^2 \leq \lambda_n \|\Pi_n \phi_1\|_k^2 \quad \text{and} \quad \|\Pi_n^\perp \phi_1\|_k^2 \leq \lambda_n^{-1} \|\Pi_n^\perp \phi_1\|_{k+1}^2,$$

we have an appropriate bound for the control term

$$\begin{aligned}(G(\mathbf{q}, \tilde{\mathbf{q}})), \phi_1) &= a(\Delta\phi_1 - \Pi_n^\perp \Delta\phi_1, \phi_1) \\ &= -ah_1\|\phi_1\|^2 + ah_1\|\Pi_n^\perp \phi_1\|^2 \\ &\leq -ah_1\|\phi_1\|^2 + ah_1\lambda_n^{-1}|\Delta\phi_1|^2. \end{aligned} \tag{69}$$

Given the estimate (69) for the control, we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||_{-1}^2 + \nu|\Delta\boldsymbol{\phi}|^2 + rh_2\|\phi_2\|^2 \leq (B(\boldsymbol{\psi}, \boldsymbol{\phi}), \boldsymbol{\phi}) - ah_1\|\phi_1\|^2 + ah_1\lambda_n^{-1}|\Delta\phi_1|^2.$$

By (21) and (19) we know that

$$|(B(\boldsymbol{\psi}, \boldsymbol{\phi}), \boldsymbol{\phi})| = |(B(\boldsymbol{\phi}, \boldsymbol{\phi}), \boldsymbol{\psi})| \leq k_0\|\boldsymbol{\phi}\||\Delta\boldsymbol{\phi}||\Delta\boldsymbol{\psi}|,$$

and by Young inequality, given $k_B = k_0^2/2\nu$ we have

$$|(B(\boldsymbol{\psi}, \boldsymbol{\phi}), \boldsymbol{\phi})| \leq \tfrac{\nu}{2}|\Delta\boldsymbol{\phi}|^2 + k_B|\Delta\boldsymbol{\psi}|^2\|\boldsymbol{\phi}\|^2.$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||_{-1}^2 + \nu|\Delta\boldsymbol{\phi}|^2 + 2ah_1\|\phi_1\|^2 + 2rh_2\|\phi_2\|^2 \leq 2k_B|\Delta\boldsymbol{\psi}|^2\|\boldsymbol{\phi}\|^2 + 2ah_1\lambda_n^{-1}|\Delta\phi_1|^2,$$

and in particular setting $a = r$,

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||_{-1}^2 + \left(\nu - 2r\lambda_n^{-1}\right)|\Delta\boldsymbol{\phi}|^2 \leq \|\boldsymbol{\phi}\|^2\left(2k_B|\Delta\boldsymbol{\psi}|^2 - 2r\right). \tag{70}$$

Choosing n so that

$$\nu - 2r\lambda_n^{-1} > 0, \tag{71}$$

Gronwall's lemma gives that

$$|||\boldsymbol{\xi}(t)|||_{-1}^2 \leq |||\boldsymbol{\xi}(0)|||_{-1}^2 \exp\left(-2rt + 2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\, ds\right).$$

Assumption **A1** follows immediately with constants $\kappa_0 = 2r$ and $\kappa_1 = 2k_B$.

Proof of A2. Consider the original model (9) and let us apply Itô formula to compute $d|||\mathbf{q}|||_{-1}^2$. Since $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$ we have

$$-d|||\mathbf{q}|||_{-1}^2 = d(\mathbf{q}, \boldsymbol{\psi}) = -d\left(\mathbf{q}, \tilde{A}^{-1}\mathbf{q}\right),$$

and since $\tilde{A}$ is self-adjoint we have

$$d\left(\mathbf{q}, \tilde{A}^{-1}\mathbf{q}\right) = 2(d\mathbf{q}, \tilde{A}^{-1}\mathbf{q}) + \mathrm{Tr}\left[(Q^{1/2})^* \tilde{A}^{-1} Q^{1/2}\right] dt.$$

Therefore, setting

$$T_Q := \mathrm{Tr}\left[(Q^{1/2})^* \tilde{A}^{-1} Q^{1/2}\right],$$

we have that $-d|||\mathbf{q}|||_{-1}^2 = 2(d\mathbf{q}, \boldsymbol{\psi}) - T_Q\, dt$, which gives

$$-d|||\mathbf{q}|||_{-1}^2 = 2\left(\nu(\Delta^2\boldsymbol{\psi}, \boldsymbol{\psi}) + h_1(f, \psi_1) - r(\Delta\psi_2, \psi_2) - \tfrac{1}{2}T_Q\right) dt + 2(\boldsymbol{\psi}, d\mathbf{W})$$

where we have used that $(B(\boldsymbol{\psi}, \boldsymbol{\psi}), \boldsymbol{\psi}) = 0$ and $(\partial_x\boldsymbol{\psi}, \boldsymbol{\psi}) = 0$. By Green's theorem and the definition of $\mathbf{W}$

$$d|||\mathbf{q}|||_{-1}^2 = -2\left(\nu|\Delta\boldsymbol{\psi}|^2 + h_1(f, \psi_1) + rh_2\|\psi_2\|^2 - \tfrac{1}{2}T_Q\right) dt - 2h_1\left(\psi_1, dW\right). \tag{72}$$

Next, using Cauchy-Schwartz, Young and Poincaré inequalities we can bound the deterministic forcing term as follows

$$-2(f, h_1\psi_1) \leq 2|(f, h_1\psi_1)| \leq \tfrac{h_1}{\nu}\|f\|_{-2}^2 + \nu h_1|\Delta\psi_1|^2 \tag{73}$$

and, using this estimate in (72), we have

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \nu\int_0^t |\Delta\boldsymbol{\psi}|^2\, ds + 2rh_2\int_0^t \|\psi_2\|^2\, ds \leq \kappa_3 t + 2h_1 X_t$$

where $\kappa_3 = \frac{h_1}{\nu}\|f\|_{-2}^2 + T_Q$ and X_t is defined by

$$X_t := \int_0^t (\psi_1(s), dW(s)).$$

The quadratic variation of this process is

$$\langle X\rangle_t := \int_0^t \|(\psi_1(s), \cdot)\|_{L_2^0}^2\, ds = \int_0^t \sum_{k\in\mathbb{N}} |(\psi_1, Q^{1/2}e_k)|^2\, ds.$$

By Cauchy-Schwartz inequality, this can be bounded by

$$\langle X\rangle_t \leq \int_0^t |\psi_1|^2 \sum_{k\in\mathbb{N}} |Q^{1/2}e_k|^2\, ds = \mathrm{Tr}\, Q\int_0^t |\psi_1(s)|^2\, ds. \tag{74}$$

Then, since X_t is a continuous martingale, it can be shown that for all $\gamma > 0$ and $R > 0$

$$\mathbb{P}\left(\sup_{t\geq 0}\left(X_t - \gamma\langle X\rangle_t\right) > R\right) \leq e^{-2\gamma R}$$

hence, setting $\Xi_\gamma := \sup_{t\geq 0}\left(X_t - \gamma\langle X\rangle_t\right)$,

$$\mathbb{E}\exp(K\Xi_\gamma) < \infty \quad \text{for all } K < 2\gamma.$$

Therefore with a simple manipulation we get

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \nu \int_0^t |\Delta \boldsymbol{\psi}|^2 \, ds + 2rh_2 \int_0^t \|\psi_2\|^2 \, ds \leq \\ \kappa_3 t + 2h_1 \Xi_\gamma + 2h_1 \gamma \langle X \rangle_t. \quad (75)$$

Using Poincaré inequality twice in (74) we obtain

$$\langle X \rangle_t \leq \frac{\operatorname{Tr} Q}{\lambda_1^2} \int_0^t |\Delta \psi_1|^2 \, ds$$

and using this in (75) gives

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \left(\nu - \tfrac{2\gamma \operatorname{Tr} Q}{\lambda_1^2}\right) \int_0^t |\Delta \boldsymbol{\psi}|^2 \, ds - \kappa_3 t \leq 2h_1 \Xi_\gamma.$$

Finally, Assumption **A2** is satisfied with

$$\kappa_2 = \nu - \tfrac{2\gamma \operatorname{Tr} Q}{\lambda_1^2} \quad \text{and} \quad \kappa_3 = \tfrac{h_1}{\nu} \|f\|_{-2}^2 + T_Q,$$

for all arbitrary parameter $\gamma > 0$ such that $\kappa_2 > 0$, and choices of parameters of the system such that $\kappa_0 > \kappa_1 \kappa_3 / \kappa_2$. For example pick $\gamma = \lambda_1^2 \nu / 4 \operatorname{Tr} Q$ so that

$$r > \tfrac{2k_B}{\nu} \left(\tfrac{h_1}{\nu} \|f\|_{-2}^2 + T_Q \right) =: r_0 \quad (76)$$

Proof of A3. Recall the generalized Poincaré inequality $|\Pi_n \Delta \varphi|^2 \leq \lambda_n \|\Pi_n \Delta \varphi\|_{-1}^2$. Then

$$|G(\mathbf{q}, \tilde{\mathbf{q}})|^2 = |a \Pi_n \Delta(\psi_1 - \tilde{\psi}_1)|^2 \leq \lambda_n a^2 \|\Pi_n \Delta(\psi_1 - \tilde{\psi}_1)\|_{-1}^2 \\ \leq \lambda_n a^2 \|\psi_1 - \tilde{\psi}_1\|^2 \leq \lambda_n a^2 |||\mathbf{q} - \tilde{\mathbf{q}}|||_{-1}^2$$

giving the desired inequality with $c = a^2 \lambda_n$.

Proof of A4. We want to show that $V(\mathbf{x}) := |||\mathbf{x}|||_{-1}^2$ satisfies Assumption **A4**. Integrating (72) over $[s, t]$ we obtain, by dropping the term $rh_1 \|\psi_2\|^2$ and estimating the forcing term as in (73),

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}(s)|||_{-1}^2 + \nu \int_s^t |\Delta \boldsymbol{\psi}|^2 \, d\tau - (t-s) T_Q \leq \\ \tfrac{h_1 \|f\|_{-2}^2}{\nu} (t-s) - 2h_1 \int_s^t (\psi_1, dW).$$

Rearranging and taking the expectation we have

$$\mathbb{E} \, |||\mathbf{q}(t)|||_{-1}^2 \leq \mathbb{E} \, |||\mathbf{q}(s)|||_{-1}^2 + \mathbb{E} \int_s^t (-\nu |\Delta \boldsymbol{\psi}|^2 + K) \, ds$$

where $K = \frac{h_1|f|^2_{-2}}{\nu} + T_Q$. By Equation (16) and Poincaré inequality we know that $|||\mathbf{q}|||^2_{-1} \leq |\Delta\boldsymbol{\psi}|^2$, so that

$$\mathbb{E}\,|||\mathbf{q}(t)|||^2_{-1} \leq \mathbb{E}\,|||\mathbf{q}(s)|||^2_{-1} + \mathbb{E}\int_s^t \left(-\tfrac{\nu\lambda_1}{a_0}|||\mathbf{q}(\tau)|||^2_{-1} + K\right)\, d\tau.$$

Therefore $V(\mathbf{q}) = |||\mathbf{q}|||^2_{-1}$ satisfies the estimate (32) with

$$\gamma_1 := \frac{\nu\lambda_1}{a_0} \quad \text{and} \quad K = \frac{h_1\|f\|^2_{-2}}{\nu} + T_Q.$$

□

Remark 4.2. The existence of an invariant measure can also be proved without conditions on the parameter r or any other parameter of the model. In fact it can be shown by means of the classic Krylov–Bogoliubov theorem, similarly to what was done for the 2D Navier–Stokes equations in [17]. For a complete proof of the existence of the invariant measure of the stochastic two–layer quasi–geostrophic model with this technique refer to [7].

Remark 4.3 (Finite dimensional noise). From the literature (e.g. [6, 19]) it is known that the coupling method applies also when the noise acts only on finitely many modes, as long as enough of them are activated. That lower bound on the dimension of the noise arose also in the argument just presented, when we required condition (ii) in Theorem 4.1. Therefore, with few modifications to the proof of Theorem 4.1, the ergodicity holds also for the model perturbed on the top layer only by a n dimensional noise as long as (71) holds.

Remark 4.4. It is interesting to notice that the generalised coupling method used provides also a description of a potential way by which the system stabilizes. Indeed the feedback control we introduced, namely $\Delta(\psi_1 - \tilde{\psi}_1)$, contains information only from the first layer. Then the condition on the bottom friction corresponds to a scenario in which the first layer stabilizes by the influence of the stochastic forcing and the second layer stabilizes mainly thanks to its friction.

Remark 4.5. The result in Theorem 4.1 holds also under conditions not necessarily involving the parameter r as in Theorem 4.1. In fact with a simple modification in the proof of we can retrieve a condition also, or solely, involving the viscosity. From (70), namely

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||^2_{-1} + \left(\nu - 2r\lambda_n^{-1}\right)|\Delta\boldsymbol{\phi}|^2 \leq \|\boldsymbol{\phi}\|^2\left(2k_B|\Delta\boldsymbol{\psi}|^2 - 2r\right),$$

where n is such that $\nu - 2a\lambda_n^{-1} > 0$, we can also use Poincaré inequality and not drop the viscosity to get

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||^2_{-1} + \lambda_1\left(\nu - 2r\lambda_n^{-1}\right)\|\boldsymbol{\phi}\|^2 \leq \|\boldsymbol{\phi}\|^2\left(2k_B|\Delta\boldsymbol{\psi}|^2 - 2r\right).$$

Using (16) i.e. $\|\boldsymbol{\phi}\|^2 \leq |||\boldsymbol{\xi}|||^2_{-1} \leq a_0\|\boldsymbol{\phi}\|^2$ we derive

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\boldsymbol{\xi}|||^2_{-1} \leq |||\boldsymbol{\xi}|||^2_{-1}\left(2k_B|\Delta\boldsymbol{\psi}|^2 - 2r - \tfrac{\lambda_1}{a_0}\left(\nu - 2r\lambda_n^{-1}\right)\right),$$

so that, thanks to Gronwall lemma,

$$|||\boldsymbol{\xi}(t)|||^2_{-1} \leq |||\boldsymbol{\xi}(0)|||^2_{-1}\exp\left(-t\left(2r + \tfrac{\lambda_1}{a_0}(\nu - 2r\lambda_n^{-1})\right) + 2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\right).$$

Then, Assumption **A1** holds with $\kappa_0 = 2r + \frac{\lambda_1}{a_0}(\nu - 2r\lambda_n^{-1})$. Since Assumption **A2** holds when $\kappa_0 > \kappa_1\kappa_3/\kappa_2$, we require

$$2r + \tfrac{\lambda_1}{a_0^2}\left(\nu - 2r\lambda_n^{-1}\right) > \tfrac{2k_B}{\nu - \frac{2\gamma\,\mathrm{Tr}\,Q}{\lambda_1^2}}\left(T_Q + \tfrac{h_1}{\nu}\|f\|^2_{-2}\right).$$

where $0 < \gamma < \lambda_1^2\nu/2\,\mathrm{Tr}\,Q$. For example then picking $\gamma = \lambda_1^2\nu/4\,\mathrm{Tr}\,Q$ we have

$$2r + \tfrac{\lambda_1}{a_0}\left(\nu - 2r\lambda_n^{-1}\right) > \tfrac{4k_B}{\nu}\left(T_Q + \tfrac{h_1}{\nu}\|f\|^2_{-2}\right). \tag{77}$$

In particular this result provides exponential ergodicity of the model also when $r = 0$ as long as the viscosity is large enough. The presence of a large viscosity would also imply that we could consider smaller values of n, namely more degenerate noise on the first layer. A similar result holds also for the stochastic Navier-Stokes equation. In fact in [26] ergodicity is ensured in a large viscosity scenario even with a finite dimensional stochastic forcing.

Related to Remark 4.5, it is clear, on the one hand, that from a physical point of view we can expect ergodicity in case there is strong dissipation on both layers, for example by means of a large viscosity. On the other hand, the imposed parameter condition (76) requires sufficient dissipation only on one of the two layers by requiring the bottom layer (the one without noise) to be enslaved by the top one, or to converge autonomously, by means of a minimum requirement for the friction. However, a natural question which arises in this context is whether the spectral gap can be shown even when no particular condition on the dissipation is imposed. This is not clear directly from our analysis nor the available literature, nor does there exists a clear physical intuition. This will be subject of future research.

References

[1] R. V. Abramov and A. J. Majda. Low-frequency climate response of quasi-geostrophic wind-driven ocean circulation. *Journal of Physical Oceanography*, 42(2):243 – 260, 2012.

[2] P. S. Berloff. Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, 529:71–95, 2005.

[3] C. Bernier. Existence of attractor for the quasi-geostrophic approximation of the Navier-Stokes equations and estimate of its dimension. *Adv. Math. Sci. Appl.*, 4(2):465–489, 1994.

[4] J. R. Brannan, J. Duan, and T. Wanner. Dissipative quasi-geostrophic dynamics under random forcing. *Journal of Mathematical Analysis and Applications*, 228(1):221 – 233, 1998.

[5] O. Butkovsky. Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.*, 24(2):526–552, 04 2014.

[6] O. Butkovsky, A. Kulik, and M. Scheutzow. Generalized couplings and ergodic rates for SPDEs and other Markov models. *Ann. Appl. Probab.*, 30(1):1–39, 02 2020.

[7] G. Carigi. *Ergodic properties and response theory for a stochastic two-layer model of geophysical fluid dynamics.* PhD thesis, University of Reading, 2021. https://centaur.reading.ac.uk/102181/.

[8] G. Carigi, T. Kuna, and J. Bröcker. Linear and fractional response for dissipative SPDEs. *In preparation.*

[9] J. G. Charney. On the scale of atmospheric motions. *Geofysiske Publikasjoner*, 17(2):1–17, 1948.

[10] K. C. Chhak, A. M. Moore, R. F. Milliff, G. Branstator, W. R. Holland, and M. Fisher. Stochastic forcing of the north atlantic wind-driven ocean circulation, part I: A diagnostic analysis of the ocean response to stochastic forcing. *Journal of Physical Oceanography*, 36(3):300 – 315, 2006.

[11] I. Chueshov, J. Duan, and B. Schmalfuss. Probabilistic dynamics of two-layer geophysical flows. *Stochastics and Dynamics*, 01(04):451–475, 2001.

[12] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions.* Cambridge University Press, 2014.

[13] T. DelSole. Stochastic models of quasigeostrophic turbulence. *Surveys in Geophysics*, 25(2):107–149, 2004.

[14] J. Duan and B. Goldys. Ergodicity of stochastically forced large scale geophysical flows. *International J. Math. Math. Sci*, pages 313–320, 2001.

[15] J. Duan, P. E. Kloeden, and B. Schmalfuss. Exponential stability of the quasigeostrophic equation under random perturbations. *Progress in Probability*, 49:241–256, 2000.

[16] R. M. Dudley. *Real Analysis and Probability.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2002.

[17] F. Flandoli. Dissipativity and invariant measures for stochastic Navier-Stokes equations. *Nonlinear Differential Equations and Applications NoDEA*, 1(4):403–423, 1994.

[18] M. Ghil and V. Lucarini. The physics of climate variability and climate change. *Rev. Mod. Phys.*, 92:035002, Jul 2020.

[19] N. Glatt-Holtz, J. C. Mattingly, and G. Richards. On unique ergodicity in nonlinear stochastic partial differential equations. *Journal of Statistical Physics*, 166(3-4):618–649, 2017.

[20] A. Griffa and S. Castellari. Nonlinear general circulation of an ocean model driven by wind with a stochastic component. *Journal of Marine Research*, 49:53–73, 02 1991.

[21] M. Hairer. Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probability Theory and Related Fields*, 124:345–380, 2001.

[22] M. Hairer and A. J. Majda. A simple framework to justify linear response theory. *Nonlinearity*, 23(4):909, 2010.

[23] M. Hairer and J. C. Mattingly. Yet another look at Harris'ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117. Springer Basel, 2011.

[24] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probability Theory and Related Fields*, 149(1):223–259, Feb 2011.

[25] A. Kulik and M. Scheutzow. Generalized couplings and convergence of transition probabilities. *Probability Theory and Related Fields*, 171(1):333–376, 2018.

[26] J. C. Mattingly. Ergodicity of 2D Navier–Stokes equations with random forcing and large viscosity. *Communications in Mathematical Physics*, 206(2):273–288, Oct 1999.

[27] J. C. Mattingly. Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. *Communications in Mathematical Physics*, 230(3):421–462, 2002.

[28] S. Meyn, R. L. Tweedie, and P. W. Glynn. *Markov Chains and Stochastic Stability*. Cambridge Mathematical Library. Cambridge University Press, 2nd edition, 2009.

[29] A. M. Moore. Wind-induced variability of ocean gyres. *Dynamics of Atmospheres and Oceans*, 29(2):335–364, 1999.

[30] C. H. O'Reilly, A. Czaja, and J. LaCasce. The emergence of zonal ocean jets under large-scale stochastic wind forcing. *Geophysical research letters*, 39(11), 2012.

[31] A. M. Treguier and B. L. Hua. Oceanic quasi-geostrophic turbulence forced by stochastic wind fluctuations. *Journal of Physical Oceanography*, 17(3):397 – 411, 1987.

[32] G. K. Vallis. *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press, Cambridge, U.K., 2006.

[33] C. Villani. *Optimal transport: old and new*. Springer, Berlin, 2008.

[34] D. Yang and J. Duan. Large deviations for the stochastic quasigeostrophic equation with multiplicative noise. *Journal of Mathematical Physics*, 51(5):053301, 2010.