Noise-induced vortex-splitting stratospheric sudden warmings

J. G. Esler*and M. Mester

Department of Mathematics, University College London

*Correspondence to: Department of Mathematics, University College London, Gower Street, London, WC1E 6BT, UK; j.g.esler@ucl.ac.uk

Observed oscillations of the Antarctic stratospheric polar vortex often resemble those in Kida's model of an elliptical vortex in a linear background flow. Here, Kida's model is used to investigate the dynamics of 'vortex splitting' stratospheric sudden warmings (SSWs), such as the Antarctic event of 2002. SSWs are identified with a bifurcation in the periodic orbits of the model. The influence of 'tropospheric macroturbulence' on the vortex is modelled by allowing the linear background forcing flow to be driven by a random process, with a finite decorrelation time (an Ornstein-Uhlenbeck process). It is shown that this stochasticity generates a random walk across the state-space of periodic orbits, which will eventually lead to the bifurcation point after which an SSW will occur. In certain asymptotic limits, the expected time before an SSW occurs can be found by solving a 'first passage time' problem for a stochastic differential equation, allowing the dependence of the expected time to an SSW on the model parameters to be elucidated. Results are verified using both Kida's model and single-layer quasi-geostrophic simulations. The results point towards a 'noise-memory' paradigm of the winter stratosphere, according to which the forcing history determines whether the vortex is quiescent, undergoes large amplitude nonlinear oscillations or, in extreme cases, whether it will split.

Key Words: sudden warmings; vortex splits; stochastic processes

Received . . .

Citation: . . .

1. Introduction

It is well-documented that stratospheric sudden warmings (SSWs hereafter) exert a significant influence on surface climate in the Northern hemisphere, following Baldwin and Dunkerton (2001) who showed that stratospheric circulation anomalies following an SSW often descend into the troposphere, where they may persist for several weeks. A similar influence can be expected in the Southern hemisphere where there has been just a single recorded SSW (2002) in the observational record (c. 1948-present). SSWs are naturally categorised into two types (e.g. Charlton and Polvani 2007): vortex displacement events, in which the vortex is displaced off the pole and eroded at upper levels, and vortex splitting events, in which the vortex divides almost simultaneously at all levels. The question of which type of SSW has a stronger influence on surface climate has been addressed by Nakagawa and Yamazaki (2006) and Mitchell *et al.* (2013), and it turns out that observations suggest that splitting events are responsible for almost all of the tropospheric response (see e.g. Fig. 4 of Mitchell *et al.* 2013). (Interestingly, however, the model results of Maycock and Hitchcock (2015) do not support this conclusion.) It is consequently of great interest to understand the fluid dynamics that determines the frequency of vortex splitting SSWs in particular, and especially how this frequency might change in a changing climate.

There have been a number of studies aimed at assessing SSW frequency under plausible scenarios for both greenhouse gas emissions and ozone recovery, using atmosphere-only mechanistic models (Butchart *et al.* 2000), chemistry-climate models (Ayarzagüena *et al.* 2013; Mitchell *et al.* 2012a) and coupled ocean-atmosphere models (Mitchell *et al.* 2012b). Overall, the results of these studies are indeterminate, with some suggestion of changes in the timing of SSWs, but no statistically significant changes in their frequency. Evidently, both computational constraints on integration length / ensemble size, and the overall complexity of global models, make it challenging to obtain a clear dynamical understanding of the processes controlling SSW frequency. A complementary approach, to be pursued below, is to study the factors controlling SSWs in a simple dynamical system, where the parameter dependencies can be fully elucidated. In particular the aim here is to investigate the effect of unsteadiness (i.e. the 'noise' of the title), caused for example by time-dependent tropospheric dynamics, in the forcing of the stratospheric vortex.

The idea that 'noise' has an important role in SSW variability has previously been investigated by Birner and Williams (2008). Using a simple model based on a dynamical reduction of the Holton-Mass model (Holton and Mass 1976; Ruzmaikan *et al.* 2003), with the noise being a stochastic forcing that models the effect of dissipating gravity waves on the stratospheric circulation, they showed how both the probability of an SSW occurring, as well as its timing, can depend on the details of the noise. Here we aim to go further than the Birner-Williams study in the following respects:

- By using a dynamical system with prognostic variables (vortex aspect ratio and orientation) that can be easily and unambiguously compared with the observed polar vortices.
- By the same dynamical system having a quantitative link to a single-layer quasi-geostrophic model which can simulate realistic-looking vortex splits.
- By demonstrating that the presence of realistic noise is, without invoking any other mechanism, sufficient to lead to winter periods with either a quiescent vortex, a vortex undergoing nonlinear oscillations in aspect ratio, or in extreme cases a split.

Of course a simple dynamical model has its limitations and, because of the chaotic nature and vertical variability of the flow in the Northern winter stratosphere, we do not claim for our model more than paradigmatic relevance to the Arctic. In the Antarctic, by contrast, our model will be

Prepared using qjrms4.cls

argued below to have relevance to observations despite its simplicity.

The simple dynamical system in question is Kida's model (Kida 1981) of an elliptical vortex patch in a linear background flow. The restriction to a two-dimensional model is justified by the near barotropic structure of observed vortex-splitting SSWs (e.g. Matthewman *et al.* 2009). In Kida's model the vortex evolves under the influence of a linear strain flow and a solid body rotation, under which conditions it remains elliptical at subsequent times, with the evolution of its aspect ratio and orientation governed by a pair of coupled differential equations (see below). The linear background flow in the model can be interpreted as a representation of the cumulative dynamical influence of the Earth's surface and the troposphere on the vortex. The idea is that, invoking 'piecewise potential vorticity inversion' (Nielsen-Gammon and Lefevre 1996), the influence of tropospheric planetary-scale stationary waves, surface topography and land-sea contrast on the vortex can (to a good approximation) be replaced by a local advecting velocity field, the 'forcing velocity'. Further, the largest-scale component of this forcing velocity, which is of the greatest dynamical significance for the vortex, can be approximated in the vicinity of the vortex by a linear flow.

Using the insights above, Matthewman and Esler (2011, ME11 hereafter) showed that Kida's equations can closely track the dynamics of a 2D quasi-geostrophic model of the stratospheric polar vortex forced by surface topography, up to the time when a vortex split is initiated in the latter model. Across much of parameter space, the elliptical vortex in Kida's model undergoes periodic nonlinear oscillations in aspect ratio and orientation. ME11 showed that vortex splits in the quasi-geostrophic model can be associated with a discrete jump in the amplitude of these oscillations, which for a given initial condition occurs across a fixed curve in parameter space. Amplitude bifurcations of exactly this type also occur in generic weakly nonlinear models of forced waves near resonance (Plumb 1981; Esler and Matthewman 2011), and in the present context the mechanism associated with the increase in Rossby wave amplitude leading to SSWs has been termed 'nonlinear self-tuning resonance'.

In the ME11 description the tropospheric forcing (linear background flow) is constant in time. In reality, the forcing experienced by the polar vortex has a significant unsteady component, due to for example propagating tropospheric planetary waves (e.g. Scinocca and Haynes 1998), and to random variability in the tropospheric circulation as a result of 'tropospheric macro-turbulence' (Held 1999). The present work will show how unsteady forcing can lead to vortex splits, both in Kida's model, and in a single layer quasi-geostrophic numerical model.

In section 2 ERA-Interim reanalysis data (Dee *et al.* 2011) is analysed to demonstrate that the stratospheric vortex in the Southern hemisphere undergoes nonlinear oscillations which share many characteristics with the oscillations of Kida's vortex. In section 3 Kida's model and its deterministic behaviour are reviewed, and mathematical results describing its behaviour under stochastic forcing are elucidated. The first passage time problem for SSWs is defined and then solved in two different asymptotic limits, and for two different types of stochastic forcing. In section 4, numerical integrations are presented which illustrate the behaviour of Kida's equations over a wide range of parameters, and the validity and relevance of the results of section 3 are explored using large-ensemble integrations. The results are compared with integrations of a 2D quasi-geostrophic model. Finally in 5 conclusions are drawn.

2. Kida-like oscillations of the Antarctic stratospheric polar vortex

Elliptical diagnostics (Waugh 1997; Waugh and Randel 1999) provide a quantitative method to describe the time-evolution of the polar vortices in terms of a few time series (see also Mitchell *et al.* 2011). Here, ERA-Interim Ertel's potential vorticity data, on the 600 K isentropic level, has been used to calculate the aspect ratio $\lambda(t)$ and the orientation $\theta(t)$ of the Antarctic vortex during the late

Prepared using qjrms4.cls

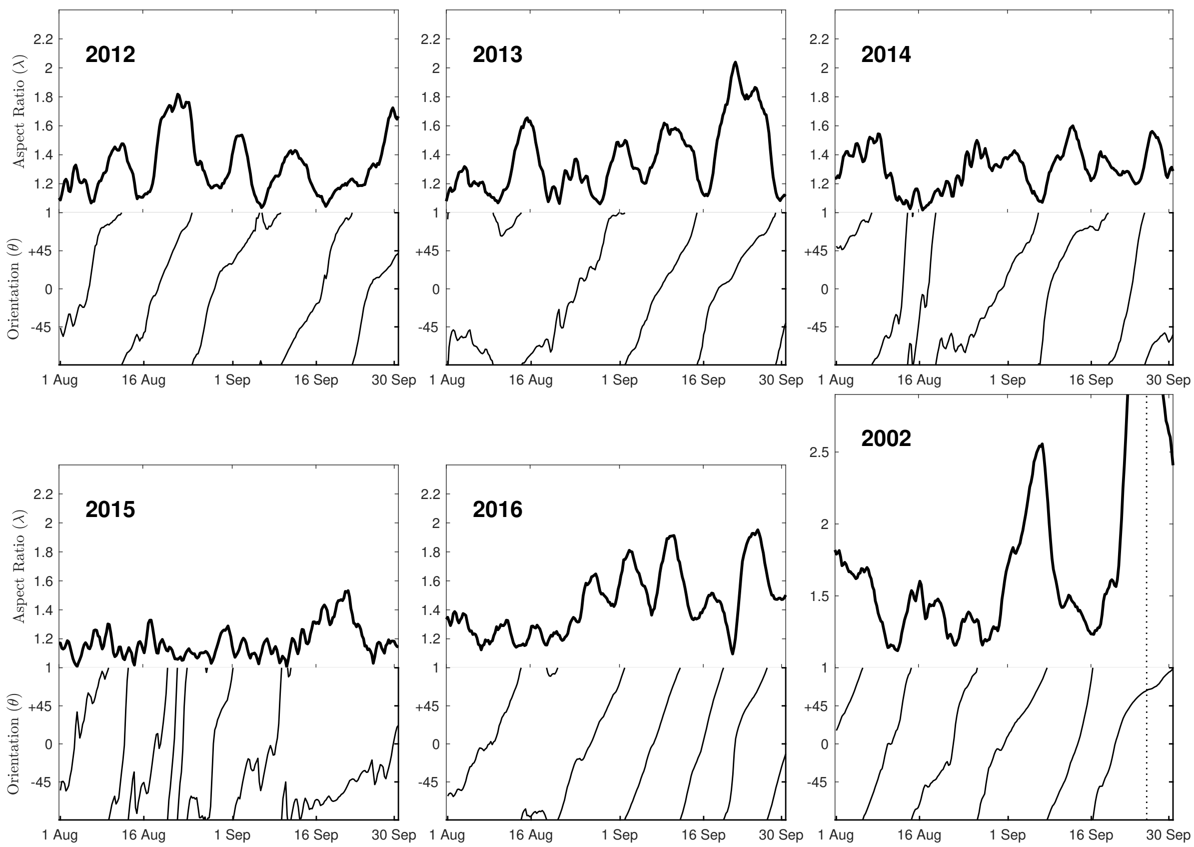

Figure 1. Antarctic polar vortex aspect ratio $\lambda(t)$ and orientation $\theta(t)$ during the late austral winters (0000UT 1 Aug- 1800UT 30 Sep) of 2012-2016 and 2002. The dashed vertical line on the 2002 panel marks the time of the Antarctic SSW measured by the WMO criterion.

austral winter (August-September) for five recent seasons (2012-2016) and for 2002 (the year of the Antarctic SSW). The procedure for calculating λ and θ from the data follows that described in section 2 of Matthewman *et al.* (2009) exactly. One technical point, however, is that θ here is measured in the same sense as longitude, which in the Southern hemisphere is in the opposite sense to the usual polar coordinates. Following this convention means that the observed results, for the negative PV Antarctic vortex, can be compared directly to the (positive vorticity) Kida vortex without further transformation. Very similar pictures emerge if other vertical levels are chosen, although it is notable that, unlike in the case of typical Arctic vortex splits the Antarctic SSW of 2002 has significant vertical structure (e.g. Esler *et al.* 2006), because at very low levels ($\sim$450 K) the vortex recovers instead of splitting. The aspect ratio in the 2002 panel in late September is therefore somewhat sensitive to the level chosen.

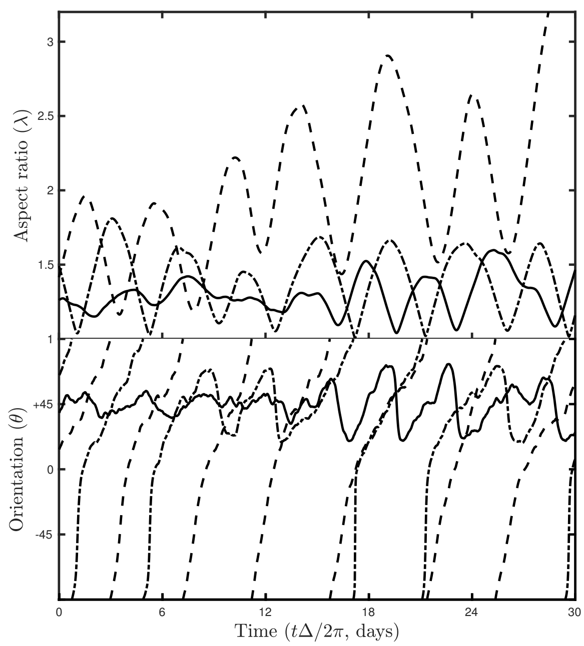

Figure 2. Sample paths from integrations of Kida's equations, with $\Gamma = 0.04$ and the rotation rate $\Omega(t)$ driven by the Ornstein-Uhlenbeck process (16), with $\Omega_0 = -0.12$, $\delta = 2\pi\Delta^{-1}$ and $\varepsilon = 0.025\delta^{-1/2}$.

Figure 1 shows $\lambda(t)$ and $\theta(t)$ for Aug-Sep 2012-2016, as well as Aug-Sep 2002. The most striking features of the time-series are:

1. In certain years, notably 2012, 2013, 2016 and 2002, there is coherent cyclonic phase propagation (i.e. $\dot{\theta} > 0$) throughout almost all of the periods shown. The mean angular frequencies for these four seasons are 0.232 (2012), 0.170 (2013), 0.271 (2016) and 0.279 (2002) radians day^{-1}.
2. Near-synchronous oscillations in aspect ratio occur, with a wide range in amplitude both within and between seasons. Scatterplots (not shown) reveal that the orientation of the vortex at maximum aspect ratio varies, but that there is a significant bias towards the direction parallel with the 40°E-140°W longitude circle. Aspect ratio fluctuations with larger amplitude appear to correlate with longer oscillation periods.
3. Occasional instances of stalling in the phase propagation (e.g. 5-8 Sep 2012, 2-11 Aug 2013, 15-22 Aug 2016), occur when the vortex has low aspect ratio.
4. In other years, such as 2014 and especially 2015, there are no coherent oscillations in aspect ratio and the coherence of the phase propagation is much reduced (note that the orientation becomes ill-defined as the aspect ratio approaches unity, which explains the rapid variations in θ).

During the 2002 oscillations, the vortex aspect ratio is correlated with oscillations in the stratospheric zonal wind at 60°S (see Figs. 2 and 6 of Scaife *et al.* 2005). Oscillations in vortex aspect ratio are therefore a plausible (partial) dynamical explanation of Scaife et al.'s 'stratospheric vacillations', because, provided the vortex remains near the pole, there will be a strong anti-correlation between the vortex aspect ratio and zonal mean wind at a fixed radius (see e.g. Esler and Scott 2005). Scaife *et al.* (2005) also reported smaller amplitude stratospheric vacillations in previous winters, notably 1995 and 1996, suggesting that the oscillations shown in Fig. 1 are a recurring feature of Southern winters over a longer period.

Figure 2 shows the evolution of $\lambda(t)$ and $\theta(t)$ during three separate integrations of Kida's model, in the presence of a linear flow which includes a relatively small stochastic component. The aim of these integrations, which are described in detail below, is to demonstrate that Kida's model with 'noise' is able to reproduce *qualitatively* the main behaviours seen in Fig. 1. The qualitative behaviour is recovered despite little attempt being made to 'fit' the parameters of Kida's model to match the observations, except to make sure that the system is initialised in the cyclonically rotating (ACW) regime described by Matthewman and Esler (2011). (The extent to which a quantitative parameter fit is possible is the subject of ongoing study.)

The remarkable feature of Fig. 2 is how different the three time-series are, given that they are realisations of the same random dynamical system. The dashed curves shows a 2002-like evolution with coherent phase propagation, and increasing amplitude leading to an SSW-like event where the aspect ratio grows to a large value. The dot-dash curves show a much lower amplitude oscillation in aspect ratio, reminiscent of the 2012, 2013 and 2016 winters, with two instances of 'phase stalling' (around $t = 45\Delta^{-1}$ and $125\Delta^{-1}$, where Δ is the vorticity difference between the vortex and the background). It is also notable that the oscillation period is slightly shorter compared with the large amplitude case. The solid curve shows no coherent oscillations until towards the end of the period, and no coherent phase-propagation. This behaviour is more typical of the 2014 and 2015 winters.

Next, the Kida system and its behaviour in the case of both deterministic and stochastic linear background flows will be studied in detail.

3. The Kida vortex system and its behaviour

3.1. *Deterministic behaviour*

The starting point for our analysis is Kida's equations (Kida 1981; Dritschel 1990) for the evolution of an elliptical

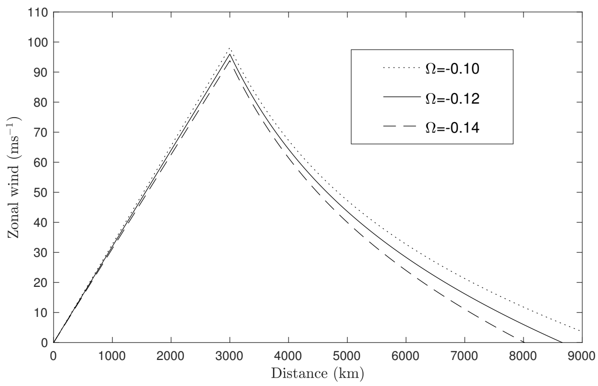

Figure 3. The ‘stratospheric’ zonal wind profile associated with the undisturbed (i.e. circular) Kida model vortex, as a function of distance from the pole, showing its sensitivity to the parameter Ω. The system is dimensionalised by choosing $\Delta = 0.5f = 2\pi$ day^{-1} and vortex radius 3000 km.

vortex patch with aspect ratio λ and orientation angle θ. The vortex evolves in a time-varying linear strain flow with amplitude $\Gamma(t)$ which is applied at angle $\Phi(t)$, and a solid body background rotation flow with rate $\Omega(t)$, according to

$$\begin{aligned} \dot{\theta} &= \Omega + \frac{\lambda}{(\lambda+1)^2} - \frac{\lambda^2+1}{\lambda^2-1}\Gamma \sin 2(\theta - \Phi) \\ \dot{\lambda} &= 2\lambda\Gamma\cos 2(\theta-\Phi). \end{aligned} \quad (1)$$

Here $\theta(t)$ is the vortex orientation and $\lambda(t)$ its aspect ratio, dots denote time derivatives, and time t, Ω and Γ are all made nondimensional using the vorticity difference Δ between the patch and the background (or its inverse), so that (1) is a nondimensional system.

Physically, following ME11, Γ can be considered to be a measure of the strength of the topographic and dynamical forcing of the vortex. The variable Ω can be associated with the current ‘climate’ in so far as it controls the zonal wind profile and its magnitude at the vortex edge. An important conceptual simplification in the model is that the forcing (i.e. $\Gamma(t)$ and $\Omega(t)$) is taken to be independent of the state of the vortex. Fig. 3 shows the stratospheric zonal wind profile induced by the undisturbed (i.e. circular) vortex, as a function of distance from the pole, illustrating the influence of Ω. Note that the cusp in the velocity at the vortex edge becomes smoothed when the vortex is elliptical or displaced slightly from the pole. Relatively small changes in Ω, which change the velocity at the vortex edge by just a few ms^{-1} will be shown below to significantly impact the expected time for an SSW. In the real atmosphere, an effective change in Ω could be caused, for example, by a change in the tropospheric Southern annular mode index. Another possibility is a change in the location of the tropical edge of the stratospheric surf zone, for example associated with the evolving quasi-biennial oscillation. In both cases, a change in the atmospheric structure away from the stratospheric vortex itself will lead, via potential vorticity inversion, to a change in the background zonal velocity at the vortex edge. Such changes can be represented in the Kida model by a change to Ω.

A key quantity for our analysis is the Hamiltonian

$$h = \frac{\lambda^2-1}{\lambda}\left(\Gamma\sin 2(\theta-\Phi) - \Omega\frac{\lambda-1}{\lambda+1}\right) - \log\frac{(\lambda+1)^2}{4\lambda}. \quad (2)$$

The physical interpretation of h, as will be explained in detail below, is that it is a quantitative measure of the character of the oscillation the vortex is undergoing. In the event that $\mathbf{\Gamma} = (\Gamma_0, \Phi_0, \Omega_0)^T$ is constant, as will be assumed throughout the present section, the system (1) can be integrated after first taking the ratio of the two equations (following Kida 1981). The result is that h is conserved by the dynamics. (As an aside, the equations (1) can be further transformed into Hamilton’s equations by transforming variables to $(p, q) = (\theta, \lambda + \lambda^{-1})$, however it does not appear to simplify the analysis below to do so). In ME11 only the case with $h = 0$ was considered. However in the situation with ‘noise’, discussed below, all values of h are accessible and so the influence of h on the nature of the oscillation must first be understood.

To understand the influence of h, square the second equation in (1), and use the definition of h to give the potential form

$$\dot{\lambda}^2 + V(\lambda) = 0, \quad (3)$$

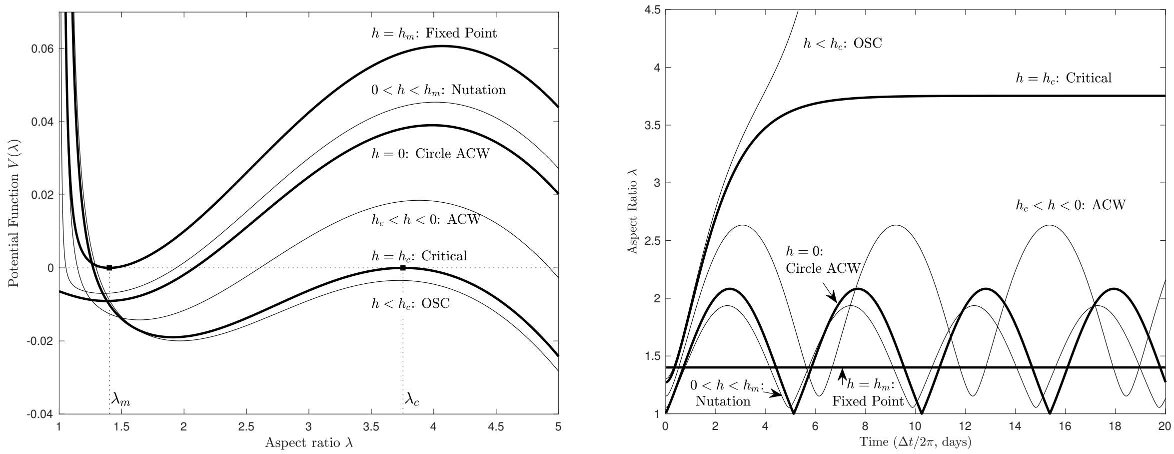

Figure 4. Left: The potential function $V(\lambda)$ for different values of h, illustrating the different regimes accessible when $(\Gamma_0, \Omega_0) = (0.04, -0.12)$. The values are $h = \{1.1h_c, h_c, 0.5h_c, 0, 0.3h_m, h_m\}$, where $h_c \approx -0.02691$ and $h_m \approx 0.01297$ are the critical and maximum values defined by (6). Right: Time evolution of the aspect ratio $\lambda(t)$, obtained from numerical integrations of (1), showing the oscillations associated with the potential functions $V(\lambda)$ in the left panel.

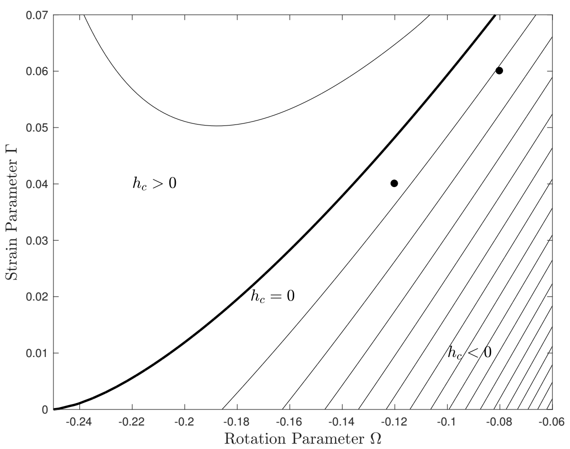

Figure 5. The critical value of h for an SSW, $h = h_c$, as a function of (Ω, Γ). The contour interval is 0.04. The zero contour, which marks the transition between ME11's ACW and OSC regimes, is marked in bold. The solid points show the parameter values used in Fig. 4, and in the simulations described in section 4.

where the potential function is defined by

$$V(\lambda) = 4\lambda^2 \left(\left(\frac{\lambda}{\lambda^2 - 1} \log \frac{e^h (\lambda + 1)^2}{4\lambda} + \Omega \frac{\lambda - 1}{\lambda + 1} \right)^2 - \Gamma^2 \right). \tag{4}$$

When the potential function satisfies $V(\lambda) < 0$ within a bounded region $\lambda_- < \lambda < \lambda_+$, i.e. a 'potential well', equation (3) is a generic equation of a nonlinear oscillator. The vortex oscillates between minimum aspect ratio λ_- and maximum λ_+, where $V(\lambda_\pm) = 0$. Further details of the nature of the oscillations depend on the structure of $V(\lambda)$ in the potential well region which can change qualitatively as h is varied. One of the key results of ME11 was to identify vortex splitting SSWs with a bifurcation associated with a qualitative change in the shape of the potential well.

Fig. 4 (left panel) shows how the shape of the potential $V(\lambda)$ changes as h is varied, with (Ω_0, Γ_0) fixed, as illustrated in Fig. 4 (left panel). Here $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$ have been chosen to fall in a region of parameter space identified by ME11 as being representative of 'typical' mid-winter stratospheric conditions (constant $\Phi_0 = 0$ is assumed without loss of generality). In ME11 a negative value of Ω_0 was found to be necessary to allow a reasonable fit to be made to the observed latitudinal profile of the stratospheric jet. It is evident that a class of relatively low amplitude ($\lambda_+ \lesssim 3.75$) oscillations of the vortex occur when h falls in the interval $h_c < h < h_m$. The upper bound $h = h_m$ corresponds to a fixed point of (1) with $\lambda = \lambda_m$ and $\theta - \Phi = \pi/4$ (the region $h > h_m$ is inaccessible). The lower bound $h = h_c$ corresponds to a critical trajectory, which reaches a maximum amplitude λ_c, and marks the SSW bifurcation identified by ME11. For $h < h_c$, a transition occurs to a regime with much larger amplitude oscillations, which is labelled OSC by ME11 and in Figs. 4-5. In the example plotted in Fig. 4 the OSC oscillation has maximum amplitude $\lambda_+ \approx 20$. Using the fact that $V(\lambda_{c,m}) = V'(\lambda_{c,m}) = 0$ it is straightforward to show that $\lambda_{c,m}$ are the two largest distinct real roots of the cubic ($\lambda_m < \lambda_c$)

$$(\Gamma_0 - \Omega_0)\lambda^3 + (\Gamma_0 - \Omega_0 - 1)\lambda^2 + (\Gamma_0 + \Omega_0 + 1)\lambda + (\Gamma_0 + \Omega_0) = 0. \quad (5)$$

It follows that the critical and maximum values of h are given by

$$h_{c,m} = (1 + 2\Omega_0)\frac{(\lambda_{c,m} - 1)^2}{\lambda_{c,m}^2 + 1} - \log\frac{(\lambda_{c,m} + 1)^2}{4\lambda_{c,m}}. \quad (6)$$

The critical value h_c is contoured as a function of (Ω_0, Γ_0) in Fig. 5. For the purposes of comparison with ME11, it is the $h_c = 0$ contour, marked in bold, which was there identified with the SSW bifurcation, because ME11 was restricted to considering the case with the initial condition taken to be a circular vortex.

At the parameter values for Fig. 4, marked with a solid point in Fig. 5, the system has $h_c < 0 < h_m$. There is therefore a further transition in the character of the oscillation at $h = 0$, between an 'anti-clockwise rotating' regime (ACW, $h_c < h < 0$) in which the major axis of the vortex rotates continuously and a 'nutating' regime ($0 < h < h_m$) in which the major axis of the vortex oscillates around the orientation $\theta - \Phi_0 = \pi/4$. The transition point between these regimes at $h = 0$ corresponds to the only trajectory to include the circular vortex ($\lambda = 1$). The time evolution of $\lambda(t)$, obtained by direct numerical integration of (1), during each type of cycle is shown in Fig. 4 (right). Note that the OSC calculation is stopped when $\lambda = 4.5$, because in both the stratosphere and in more realistic models (see below), the vortex will be unstable to perturbations at large aspect ratios, i.e. an SSW will follow once this aspect ratio is attained.

It is useful for the analysis below to introduce at this point the concept of a cycle average. Let $f(\lambda)$ be any function of aspect ratio. Its cycle average is defined to be

$$\langle f \rangle = \frac{1}{T_p}\oint_C \frac{f(\lambda)}{(-V(\lambda))^{1/2}}\,\mathrm{d}\lambda, \quad (7)$$
$$\text{where } T_p = \oint_C \frac{\mathrm{d}\lambda}{(-V(\lambda))^{1/2}}$$

is the oscillation period, obtained by direct integration of (3). The integral $\oint_C$ corresponds to integrating over a single oscillation. The integration contour C picks up the positive branch of the square root outwards along the real interval (λ_-, λ_+) and the negative branch backwards along the same interval, i.e. C should be interpreted as a *clockwise* closed contour in the complex-plane encircling (infinitesimally closely) the branch cut of $(-V(\lambda))^{1/2}$ which lies along the real axis between λ_- and λ_+. For analytic functions f it follows that $\oint_C \equiv 2\int_{\lambda_-}^{\lambda_+}$. Finally, it will also be helpful to introduce the *cycle variance* $\langle\langle f \rangle\rangle$, which is defined to be

$$\langle\langle f \rangle\rangle = \langle (f - \langle f \rangle)^2 \rangle. \quad (8)$$

3.2. *Stochastic behaviour*

A simple way of introducing the effects of 'tropospheric macroturbulence' into Kida's model is to allow the parameters $\boldsymbol{\Gamma} = (\Gamma, \Phi, \Omega)^T$ in (1) to evolve in time, and to be driven by stochastic processes. Below, the main cases that will be considered are when Γ and Ω are driven by Ornstein-Uhlenbeck processes. However, it is helpful for the analysis to first consider a rather more general possibility

$$\mathrm{d}\boldsymbol{\Gamma} = \varepsilon_* \boldsymbol{F}(\boldsymbol{\Gamma})\mathrm{d}t + \varepsilon_*^{1/2}\boldsymbol{\Sigma}(\boldsymbol{\Gamma}) \cdot \mathrm{d}\boldsymbol{W}, \quad (9)$$

where $\boldsymbol{W} = (W_1, W_2, W_3)^T$ is a three-dimensional Brownian (Wiener) process and ε_* is a nondimensional parameter introduced as a measure of the strength of the noise and drift. Here $\boldsymbol{F} = (F^\Gamma, F^\Phi, F^\Omega)^T$ is a general vector-valued 'drift' and $\boldsymbol{\Sigma}$ a 'noise' matrix, which for simplicity we take below to be diagonal, i.e. $\boldsymbol{\Sigma} = \mathrm{diag}(\Sigma^\Gamma, \Sigma^\Phi, \Sigma^\Omega)$.

To facilitate our analysis, it is helpful to consider (λ, h) as the dependent variables in place of (λ, θ). Henceforth we will use capitals (Λ, H) in recognition of the fact that they are now stochastic variables. The equation for Λ is (3) in stochastic notation is

$$\mathrm{d}\Lambda = \left(-V(\Lambda, H, \boldsymbol{\Gamma})\right)^{1/2} \mathrm{d}t. \qquad (10)$$

The equation for H is obtained by applying Itô's lemma to (2), resulting in

$$\begin{aligned}\mathrm{d}H = \varepsilon_* &\left(\left(\frac{F^\Gamma}{\Gamma} - 2(\Sigma^\Phi)^2 \right) G^\Gamma - F^\Phi G^\Phi - F^\Omega G^\Omega \right) \mathrm{d}t \\ &+ \varepsilon_*^{1/2} \left(\frac{\Sigma^\Gamma}{\Gamma} G^\Gamma \mathrm{d}W_1 - \Sigma^\Phi G^\Phi \mathrm{d}W_2 - \Sigma^\Omega G^\Omega \mathrm{d}W_3 \right),\end{aligned} \qquad (11)$$

where

$$\begin{aligned} G^\Gamma &= \Omega \frac{(\lambda-1)^2}{\lambda} + \log \frac{\mathrm{e}^H(\lambda+1)^2}{4\lambda} \\ G^\Phi &= \frac{(\Lambda^2-1)\left(-V(\Lambda, H, \boldsymbol{\Gamma})\right)^{1/2}}{\Lambda^2} \\ G^\Omega &= \frac{(\lambda-1)^2}{\lambda}. \end{aligned} \qquad (12)$$

Note that some care is needed in the interpretation of (10-11), because the branch of the square root to be taken in both equations alternates with the phase of the cycle. However, as will be described below, the great advantage of using H as a prognostic variable is that, in certain limits the long-time evolution of the vortex is completely described by an averaged H-equation, with the criterion for an SSW being simply $H < h_c$.

3.3. *The cycle-averaged equation*

In the limit $\varepsilon_* \ll 1$, it is evident from (9-11) that changes in $\boldsymbol{\Gamma}$ and H over an order unity time period, such as the period T_p of an oscillation of the vortex, will be $O(\varepsilon^{1/2})$. This observation motivates the use of the method of multiple time-scales as a method for simplifying (9-11). The aim of the analysis is to obtain an equation for the evolution of H that is valid on a time-scale $\tau \gg T_p$.

Examination of (9) suggests that the new time-scale $\tau = \varepsilon_* t$, and it follows that a Wiener process $\boldsymbol{B} = \varepsilon_*^{1/2} \boldsymbol{W}$ can be defined with respect to τ, so that (9) becomes

$$\mathrm{d}\boldsymbol{\Gamma} = \boldsymbol{F}(\boldsymbol{\Gamma})\mathrm{d}\tau + \boldsymbol{\Sigma}(\boldsymbol{\Gamma}) \cdot \mathrm{d}\boldsymbol{B}, \qquad (13)$$

where $\boldsymbol{B} = (B_1, B_2, B_3)^T$. The method of multiple time-scales can now be applied to obtain an equation for the evolution of H that can be coupled with (13). The number of dependent variables in the system is thereby reduced by one.

Care is needed in implementing the method of multiple-scales in a stochastic setting, because Wiener processes naturally include variability on all time-scales. The most straightforward method is to use standard techniques (e.g. §3.4.1 of Gardiner 2009) to transform into the deterministic setting of the Fokker-Planck equation (FPE hereafter) and then apply the method of multiple-scales method to the deterministic FPE, before transforming back again. This is the approach adopted in Appendix A.

The result is the cycle-averaged equation

$$\begin{aligned}\mathrm{d}H = &\left((F^\Gamma/\Gamma - 2(\Sigma^\Phi)^2)\langle G^\Gamma \rangle - F^\Omega \langle G^\Omega \rangle \right) \mathrm{d}\tau \\ &+ \left((\Sigma^\Gamma/\Gamma)\langle G^\Gamma \rangle \mathrm{d}B_1 - \Sigma^\Omega \langle G^\Omega \rangle \mathrm{d}B_3 \right) \\ &\left((\Sigma^\Gamma/\Gamma)^2 \langle\langle G^\Gamma \rangle\rangle + (\Sigma^\Phi)^2 \langle\langle G^\Phi \rangle\rangle \right. \\ &\left. + (\Sigma^\Omega)^2 \langle\langle G^\Omega \rangle\rangle \right)^{1/2} \mathrm{d}B. \end{aligned} \qquad (14)$$

where B is a new Brownian process which is independent of $\boldsymbol{B} = (B_1, B_2, B_3)^T$. The new Brownian process B accounts for the *intra-cycle* variability of the original Brownian processes, which would otherwise be absent from the cycle-averaged equations, and is dependent on the cycle variance $\langle\langle \cdot \rangle\rangle$ of the functions G^Γ, G^Φ etc. Notice that all cycle average and cycle variance quantities are functions of (H, Γ, Ω), and that $\langle G^\Phi \rangle = 0$ due to the presence of the branch cut in G^Φ.

It is interesting to note that even the deterministic version of (14) can be useful for understanding numerical simulations of vortex splitting SSWs. For example, Liu and Scott (2015) used a global shallow water model to simulate vortex splits, in a similar set-up to ME11. One key difference was that, due to numerical stability considerations, the topographic forcing in their experiments was introduced smoothly using a linear ramp in time (i.e. in the present notation $\Gamma = \varepsilon_* F^\Gamma t$, for F^Γ constant). Equation (14) is then

$$\frac{\mathrm{d}h}{\mathrm{d}\tau} = (F^\Gamma/\Gamma)\langle G^\Gamma \rangle. \tag{15}$$

It turns out that in the relevant parameter regime $\langle G^\Gamma \rangle > 0$ (also $\langle G^\Gamma \rangle \sim \Gamma$ for $\Gamma \ll 1$), which means that the growing topography causes h to slowly increase, pushing the vortex into the $h > 0$ nutating regime as observed in the simulations (in Figs. 6 and 7 of Liu and Scott 2015, notice that the orientation oscillates about a fixed value). It is notable that the onset of vortex splitting is less abrupt in the nutating regime compared to the ACW regime ($h < 0$). In any case, the important point is that the vortex behaviour is strongly influenced by the *history* of the forcing, as it will be in the experiments to be described below.

The cycle-averaged equation (14) will be next be used to help to obtain simplified equations for the long-time dynamics of the vortex when the linear background flow is driven by Ornstein-Uhlenbeck processes.

3.4. *Forcing by Ornstein-Uhlenbeck processes*

A relevant stochastic forcing for the linear background flow in (10-11) is the Ornstein-Uhlenbeck (O-U) process. The O-U process is of interest because it is perhaps the simplest continuous random process that, unlike the Brownian or Wiener process, can be used to model a process with a finite decorrelation time (Gardiner 2009). To focus attention on a tractable problem, we will consider O-U processes driving either the rotation Ω,

$$\mathrm{d}\Omega = -\left(\frac{\Omega - \Omega_0}{\delta}\right)\mathrm{d}t + \left(\frac{2\varepsilon^2}{\delta}\right)^{1/2}\mathrm{d}W_3, \tag{16}$$

or the strain Γ,

$$\mathrm{d}\Gamma = -\left(\frac{\Gamma - \Gamma_0}{\delta}\right)\mathrm{d}t + \left(\frac{2\varepsilon^2}{\delta}\right)^{1/2}\mathrm{d}W_1. \tag{17}$$

Here Γ_0 and Ω_0 are the prescribed long-time mean values of Γ and Ω, and the W_i are Brownian processes as in (9). The parameters ε and δ are the standard deviations and decorrelation times of the O-U processes respectively.

It turns out that there are two distinct asymptotic limits in which the stochastic Kida equations, driven by either (16) or (17), can be simplified to allow analytical progress. Both limits involves using the method of multiple time-scales to obtain a cycle averaged equation, and using the method of homogenisation (e.g. Pavliotis and Stuart 2007), to average over the time-scale of the O-U process (or 'homogenise' the system on this time-scale). However, the order in which these two methods are used is different in each case.

The first limit is the 'rapid fluctuation' limit $\delta \ll 1, \varepsilon^2\delta \ll 1$. In this limit the timescale δ for the O-U process is much shorter than the oscillation period T_p, so we can treat the O-U processes as 'fast' processes which can be averaged over, using the method of homogenisation before applying cycle-averaging. The second limit is the 'slow evolution' limit, for which $\varepsilon \sim 1 \ll \delta$, and in this case the cycle-averaging can be used as the first step, followed by homogenisation. Interestingly, in both limits, the behaviour of the system is entirely governed by a random walk with drift in H.

Next, each limit is considered in turn, treating the rotation and strain O-U processes separately.

3.4.1. *Rotation O-U process: Rapid fluctuation limit*

To treat the rapid fluctuation limit, in which the decorrelation time-scale of the O-U process satisfies $\delta \ll 1$, the homogenisation method detailed in Appendix B is first applied to the system consisting of Kida's equations (1), coupled to the rotation (Ω) O-U process (16), with $\Gamma = \Gamma_0$ constant. The resulting homogenised system is

$$\begin{aligned}
\mathrm{d}\Theta &= \left(\Omega_0 + \frac{\Lambda}{(\Lambda+1)^2} - \frac{\Lambda^2+1}{\Lambda^2-1}\Gamma_0 \sin 2\Theta\right)\,\mathrm{d}t \\
&\quad + 2^{1/2}\kappa^{1/2}\mathrm{d}W \\
\mathrm{d}\Lambda &= 2\Lambda\Gamma_0 \cos 2\Theta\,\mathrm{d}t,
\end{aligned} \tag{18}$$

where $\kappa = \varepsilon^2\delta \ll 1$. To now apply cycle-averaging, notice that equation (18) can be recast into a form consistent with that in section 3.3, by writing $\varepsilon_* = \kappa$ and substituting $\bar{\Theta} = \Theta + \Phi$, where $\Phi = 2^{1/2}\varepsilon_*^{1/2}W$ to obtain

$$\begin{aligned}
\mathrm{d}\bar{\Theta} &= \left(\Omega_0 + \frac{\Lambda}{(\Lambda+1)^2} - \frac{\Lambda^2+1}{\Lambda^2-1}\Gamma_0 \sin 2(\bar{\Theta}-\Phi)\right)\,\mathrm{d}t \\
\mathrm{d}\Lambda &= 2\Lambda\Gamma_0 \cos 2(\bar{\Theta}-\Phi)\,\mathrm{d}t, \\
\mathrm{d}\Phi &= 2^{1/2}\varepsilon_*^{1/2}\mathrm{dW}.
\end{aligned} \tag{19}$$

It follows from section 3.3 that, after substituting back for the original timescales, the cycle-averaged equation is

$$\mathrm{d}H = -4\varepsilon^2\delta\langle G^\Gamma\rangle_0\mathrm{d}t + 2^{1/2}\varepsilon\delta^{1/2}\langle\langle G^\Phi\rangle\rangle_0^{1/2}\mathrm{d}W, \tag{20}$$

where the zero subscripts denote that the cycle-averages and variances are taken at the constant values (Γ_0, Ω_0), so that $\langle G^\Gamma\rangle_0$ and $\langle\langle G^\Phi\rangle\rangle_0$ are functions only of H.

The important point about (20) is that it is a stochastic differential equation in the single variable H. The drift and diffusion functions which appear are just the cycle-averages and variances of the functions in (12), which, although they can't be explicitly obtained analytically, are easily evaluated numerically when required. The criterion for the bifurcation in the closed orbits of the system (i.e. an SSW) is simply $H = h_c$. The key question of how long it will take before an SSW occurs has been reduced to the question of how long (on average) it takes for H to first reach h_c in (20). The solution to this problem will be addressed in section 3.5 below.

3.4.2. *Rotation O-U process: Slow evolution limit*

Next, the slow evolution limit ($\delta \gg 1, \varepsilon \ll 1$) is considered. In this case, the O-U forcing is already in the form (9), provided we identify ε_* with δ^{-1}. As a consequence, the the cycle-averaged equation for H, derived in section 3.3, together with the equation for Ω, written in the slow time-variables $\tau = \delta^{-1}t$ and $B_3 = \delta^{-1/2}W_3$, can be written down as

$$\begin{aligned}
\mathrm{d}\Omega &= -(\Omega - \Omega_0)\mathrm{d}\tau + 2^{1/2}\varepsilon\mathrm{d}B_3, \\
\mathrm{d}H &= (\Omega - \Omega_0)\langle G^\Omega\rangle\mathrm{d}\tau - 2^{1/2}\varepsilon\langle G^\Omega\rangle\mathrm{d}B_3 \\
&\quad + 2^{1/2}\varepsilon\langle\langle G^\Omega\rangle\rangle^{1/2}\mathrm{d}B,
\end{aligned} \tag{21}$$

where G^Ω is defined in (12) and B is an independent Wiener process in τ. Notice that at this stage the cycle-averaged quantities $\langle G^\Omega\rangle$ etc. are functions of (H, Γ_0, Ω).

Exploiting the fact that $\varepsilon \ll 1$, the system (21) can be now be homogenised to give the behaviour on time-scales much greater than τ. Following the procedure set out in Appendix B, taking care to Taylor expand functions of ω where necessary, results in

$$\begin{aligned}
\mathrm{d}H &= \varepsilon^2\delta^{-1}\left(\partial_\omega\langle G^\Omega\rangle\big|_0 - \langle G^\Omega\rangle_0\partial_h\langle G^\Omega\rangle_0\right)\mathrm{d}t \\
&\quad + 2^{1/2}\varepsilon\delta^{-1/2}\langle\langle G_0^\Omega\rangle\rangle_0^{1/2}\mathrm{d}W,
\end{aligned} \tag{22}$$

where $\partial_\omega\,\langle G^\Omega\rangle\big|_0 \equiv (\partial\langle G^\Omega\rangle/\partial\Omega)(H, \Gamma_0, \Omega_0)$, and $\partial_h\langle G^\Omega\rangle_0 \equiv (\partial\langle G^\Omega\rangle_0/\partial H)(H)$. Note that, as for (20), we have substituted back the original time-scale.

Interestingly, it turns out that (22), like the rapid fluctuation equation (20), is also a stochastic differential equation in H , albeit with rather different drift and diffusion functions. The dependence of the governing time-scale on the O-U parameters is different, here the time-scale $\sim \varepsilon^{-2}\delta$, as opposed to $\sim \varepsilon^{-2}\delta^{-1}$ in the rapid-fluctuation limit.

3.4.3. *Strain O-U process: Rapid fluctuation limit*

In order to compare the relative importance of noise in the strain component versus the rotation component of the linear background flow, we next consider Kida's equations (1), coupled to the strain (Γ) O-U process (17), this time taking $\Omega = \Omega_0$ constant.

In the rapid fluctuation limit, the short time-scale of the O-U process ($\delta \ll 1$) means that homogenisation can be used, following Appendix B, to obtain,

$$\begin{aligned}
\mathrm{d}\Theta = &\left(\Omega_0 + \frac{\Lambda}{(\Lambda+1)^2} - \frac{\Lambda^2+1}{\Lambda^2-1}\Gamma_0 \sin 2\Theta \right.\\
&\left. + \kappa \sin 4\Theta \left(\frac{(\Lambda^2+1)^2+4\Lambda^2}{(\Lambda^2-1)^2}\right)\right)\,\mathrm{d}t\\
&- 2^{1/2}\kappa^{1/2}\sin 2\Theta \left(\frac{\Lambda^2+1}{\Lambda^2-1}\right)\mathrm{d}W\\
\mathrm{d}\Lambda = &\left(2\Lambda\Gamma_0\cos 2\Theta + 4\kappa\Lambda\left(1+\frac{2}{\Lambda^2-1}\sin^2 2\Theta\right)\right)\,\mathrm{d}t\\
&+ 2^{3/2}\kappa^{1/2}\Lambda\cos 2\Theta\,\mathrm{d}W, \qquad (23)
\end{aligned}$$

where $\kappa = \epsilon^2\delta$ and W is a single Brownian process. Applying Itô's lemma (the 'chain rule' of stochastic calculus Gardiner 2009) to H then gives

$$\mathrm{d}H = -4\kappa G_0 \mathrm{d}t - 2^{1/2}\kappa^{1/2}R_0^{1/2}\mathrm{d}W \qquad (24)$$

where G_0 and R_0 are functions of λ and H given by

$$\begin{aligned}
G_0 = &\frac{2\lambda}{(\lambda+1)^2} + \Omega_0\frac{\lambda^2+1}{\lambda} \qquad (25)\\
&+ \frac{1}{\Gamma_0^2}\frac{\lambda^2}{(\lambda+1)^4}\left(\Omega_0\frac{(\lambda-1)^2}{\lambda} + \log\frac{\mathrm{e}^H(\lambda+1)^2}{4\lambda}\right)^2\\
R_0 = &-\frac{V(\lambda,H,\Gamma_0,\Omega_0)}{\Gamma_0^2\lambda^2}\left(\frac{\lambda-1}{\lambda+1} + \Omega_0\frac{\lambda^2-1}{\lambda}\right)^2. \qquad (26)
\end{aligned}$$

Applying the cycle-averaging procedure described in Appendix A results, straightforwardly, in

$$\mathrm{d}H = -4\varepsilon^2\delta\langle G_0\rangle\,\mathrm{d}t + 2^{1/2}\varepsilon\delta^{1/2}\langle R_0\rangle^{1/2}\,\mathrm{d}W_*, \qquad (27)$$

where W_* is a new Wiener process. Equation (27) is the analogue of (20) when the O-U noise is applied to the strain rather than the rotation component of the linear background flow.

3.4.4. *Strain O-U process: Slow evolution limit*

The treatment for the slow evolution limit ($\delta \gg 1$, $\varepsilon \ll 1$) for the strain O-U process is almost identical to the rotation case above. First, identifying δ^{-1} with ε_*, the cycle-averaging procedure of section 3.3 is used, to re-scale the O-U process and write down equation (14) for the evolution of H on the slow time-scale $\tau = \delta^{-1}t$ as

$$\begin{aligned}
\mathrm{d}\Gamma &= -(\Gamma-\Gamma_0)\mathrm{d}\tau + 2^{1/2}\varepsilon\mathrm{d}B_1,\\
\mathrm{d}H &= -\frac{\Gamma-\Gamma_0}{\Gamma}\langle G^\Gamma\rangle\mathrm{d}\tau + 2^{1/2}\frac{\varepsilon}{\Gamma}\langle G^\Gamma\rangle\mathrm{d}B_1\\
&\quad + 2^{1/2}\frac{\varepsilon}{\Gamma}\langle\langle G^\Gamma\rangle\rangle^{1/2}\mathrm{d}B, \qquad (28)
\end{aligned}$$

where G^Γ is defined in (12), $B_3 = \delta^{-1/2}W_3$, and B is an independent Wiener process in τ.

Equation (28) can be homogenised in an almost identical fashion to (21) (see Appendix B) giving

$$\begin{aligned}
\mathrm{d}H = &-\varepsilon^2\delta^{-1}\left(\frac{\partial_\gamma\langle G^\Gamma\rangle|_0}{\Gamma_0} - \frac{\langle G^\Gamma\rangle_0}{\Gamma_0^2} + \frac{\partial_h\langle G^\Gamma\rangle_0\langle G^\Gamma\rangle_0}{\Gamma_0^2}\right)\mathrm{d}t\\
&+ 2^{1/2}\frac{\varepsilon^{1/2}\delta^{-1}}{\Gamma_0}\langle\langle G^\Gamma\rangle\rangle_0^{1/2}\mathrm{d}W. \qquad (29)
\end{aligned}$$

Equation (29) is the analogue of (22) for the strain O-U process.

3.5. *The first passage time problem*

Next, we address the issue of how the results above can be used to gain insight into the statistics of SSWs in the model. The idea is to formulate the *first passage time problem* for the criterion for the onset of an SSW, which is then solved to obtain the expected time until an SSW occurs. Discovering how the expected SSW time depends on the model parameters then throws light on how climatic changes may affect SSW frequency in a more realistic setting.

The analysis of sections 3.4.1-3.4.4 leads, in each example, to a one-dimensional 'random walk with drift' equation for H, of the form

$$\mathrm{d}H = a(H)\,\mathrm{d}t + b(H)^{1/2}\,\mathrm{d}W. \tag{30}$$

The smooth functions $a(h)$ and $b(h) \geq 0$ in each case have an implicit dependence on the parameters $\{\Gamma_0, \Omega_0\}$, through the cycle-averaging operation. By contrast, the dependence of $a(h)$ and $b(h)$ on the parameters ε and δ of the O-U process is relatively simple in both limits, as seen above. In the first passage time problem for systems such as (30), the aim is to calculate the expected time $T(h)$ for the system to evolve from an initial condition $H(0) = h$ to meet for the first time a specific criterion. In the present case, the relevant criterion is $H = h_c$ which, based on the discussion above, will lead to the bifurcation in the vortex oscillation associated with an SSW. The first passage time time $T(h)$ is then the expected time for an SSW event to occur.

In Appendix C it shown (following e.g. section 5.2.7 of Gardiner 2009) that $T(h)$ satisfies the ordinary differential equation

$$a(h)T'(h) + \tfrac{1}{2}b(h)T''(h) = -1, \tag{31}$$

with boundary conditions

$$T(h_c) = 0, \quad T'(h_m) = 0. \tag{32}$$

The boundary value problem (31-32) has explicit solution

$$T(h) = \int_{h_c}^{h} \frac{1}{\mu(s)} \left(\int_{s}^{h_m} \frac{2\mu(q)}{b(q)}\,\mathrm{d}q \right) \mathrm{d}s, \tag{33}$$

where

$$\mu(h) = \int_{h_c}^{h} \exp\left(\frac{2a(q)}{b(q)}\right) \mathrm{d}q. \tag{34}$$

Equation (33) allows the expected time to an SSW (specifically, $T(0)$ for a circular vortex initial condition) to be calculated, provided the functions $a(h)$ and $b(h)$ can be calculated. It is not necessary to calculate a and b explicitly to obtain the dependence on the O-U parameters ε and δ. Direct insertion of the formulae above into (33) reveals that $T(0) \sim \varepsilon^{-2}\delta^{-1}$ in the rapid fluctuation limit and $T(0) \sim \varepsilon^{-2}\delta$ in the slow evolution limit. To determine the dependency on the other parameters, standard numerical quadrature is used to obtain $a(h)$ and $b(h)$ on a suitable h-grid for each of the four examples above. The result is that the dependence of the SSW time on the model parameters can be systematically calculated and explored, and the sensitivity of the system to changes in the parameters can be evaluated, as will be seen next.

4. Results

In this section numerical integrations of Kida's equations will be used to illustrate how the expected time to an SSW depends upon the model parameters. First, the regime with $\delta \sim T_p$ (i.e. the decorrelation time of the forcing is comparable to the oscillation period), in which neither asymptotic theory described above is valid, will be explored numerically. Then, the validity and practical relevance of the asymptotic results will be verified by comparing the results of numerical simulations with those calculated from the asymptotic formulae using (33). Finally, the relevance of the stochastic Kida model will be illustrated by making a careful comparison between the Kida equation simulations and a quasi-geostrophic model that simulates realistic-looking vortex splits.

4.1. Numerical integrations of Kida's equations and general model behaviour

The first main questions to be addressed concern the sensitivity of the SSW frequency to changes in the amplitude ε and decorrelation time δ of the stochastic processes forcing the system. The results are shown in Figs 6 and 7. In each of these figures, results from ensembles of $10^3 - 10^4$ simulations of (1) forced by either rotation (16) or strain (17) processes are presented. Each simulation in each ensemble is continued until the SSW time T_λ, defined to be the first time that an aspect ratio criterion $\lambda > \lambda_c$ (see below) is reached. The mean SSW time is

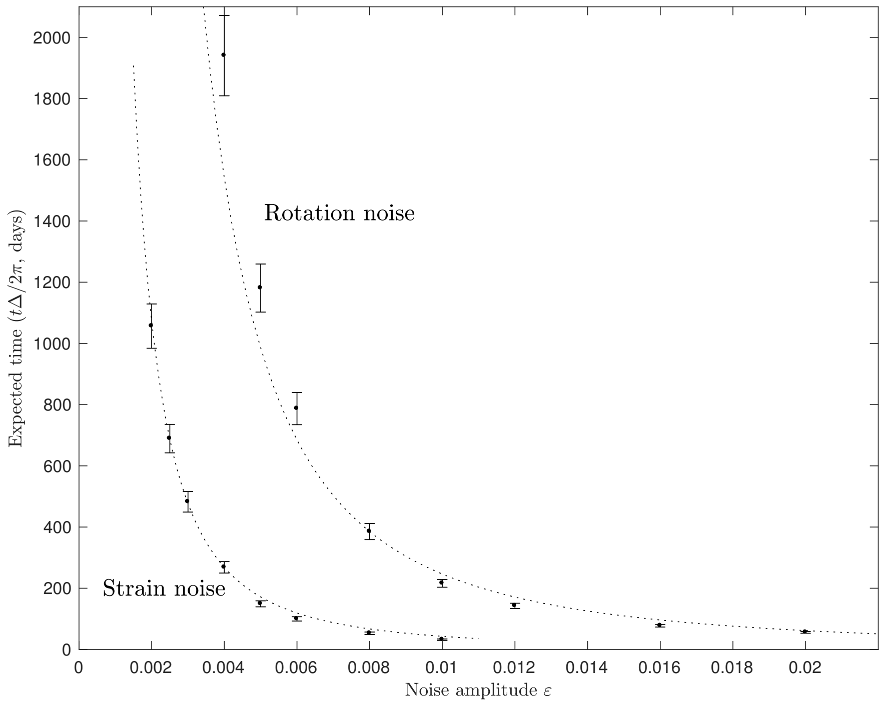

Figure 6. Expected SSW time T_λ (first time for which $\lambda > 4.5$) as a function of noise amplitude ε. Error bars show 95% confidence limits calculated from an ensemble size of 10^3. Results are plotted for both the rotation and strain O-U processes with parameters in both cases $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$ and O-U decorrelation time $\delta = 6$ 'days'.

then calculated from the ensemble and, where possible, compared with theoretical results calculated from (33).

Fig. 6 shows the mean SSW time T_λ as a function of noise amplitude ε. Results are plotted for both the rotation and strain O-U processes with parameters in both cases $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$ and a 'realistic' O-U decorrelation time $\delta = 6$'days'. Both the rapid fluctuation and slow evolution theories (valid for small and large δ respectively) predict that the SSW time should scale as ε^{-2}, and the dotted lines show 'fits' to the numerical results $\propto \varepsilon^{-2}$. The ε^{-2} scaling is a good fit in the case of the strain O-U process, but less good for the rotation O-U process, which (from a log-log fit) has a scaling closer to $\varepsilon^{-2.2}$. The numerical results therefore support the conclusions from the mathematical analysis that T_λ is sensitive to noise amplitude, which indicates that SSW frequency could be significantly affected by an increase in e.g. storm track activity associated with planetary wave generation (Scinocca and Haynes 1998).

Fig. 7 shows the mean SSW time T_λ as a function of δ, for the rotation O-U process with $\varepsilon = 0.005$ and the strain O-U process with $\varepsilon = 0.0025$. In both cases $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$. Both the rapid fluctuation theory (solid curves $\sim \delta^{-1}$) and slow evolution theory (dashed lines $\sim \delta$) are plotted against the simulation results for T_λ. The coefficients for these curves have been calculated using (33). Comparing the rotation and strain O-U processes at $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$ in Fig. 7, the mean SSW times are roughly comparable. The noise amplitude ε in the strain case is half that in the rotation case, indicating that noise on the strain component of the forcing is more than twice as effective in causing an SSW. In both cases the O-U process is most efficient in causing an SSW when $\delta \approx 1$'day', which is considerably less than a typical value of T_p of around 6 'days'.

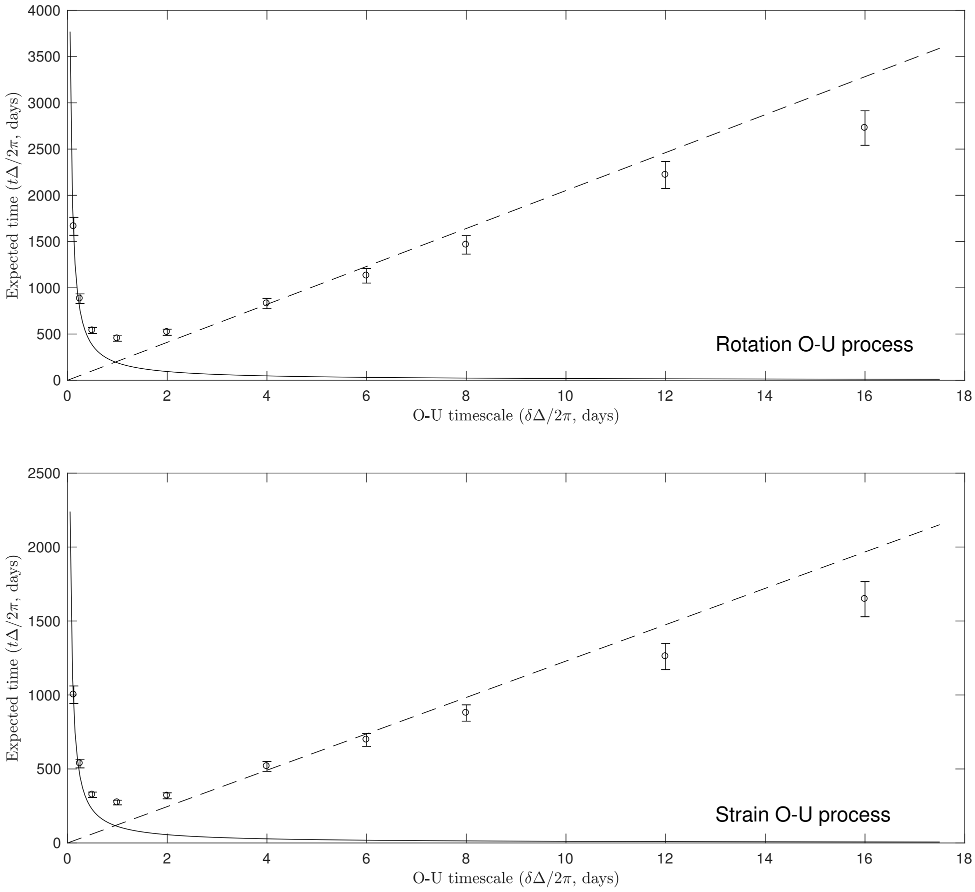

Figure 7. Mean first SSW time T_λ (points, error bars show 95% confidence limits) from simulations of Kida's equations, as a function of the O-U process timescale δ. Also plotted are the mean first passage time T_h from the rapid fluctuation theory (solid curves), and the slow evolution theory (dashed lines), each calculated using (33). Top panel: For the rotation O-U process, with amplitude $\varepsilon = 0.005$. Bottom panel: for the strain O-U process with amplitude $\varepsilon = 0.0025$. In both cases $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$.

4.2. Calculation and validation of the first-passage time formulae

Next, the dependence of the mean SSW time on the 'climate' parameters (Ω_0, Γ_0) will be elucidated. To understand the sensitivity to these parameters, it is helpful to recall Fig. 5, which shows h_c, the critical value for the Hamiltonian H, as a function of (Ω_0, Γ_0). Loosely speaking, the further away from zero is the value of h_c, the longer the system will need to reach $H = h_c$ and cause an SSW. By contrast, rapid onset of SSWs will occur for parameter settings close to the $h_c = 0$ curve on Fig. 4.

The accuracy and relevance of the asymptotic results, described in sections 3.4.1-3.5, which are formally valid only in the relevant asymptotic limits, are also tested here be at finite ε and δ. It is useful in this context to define the time T_h to be the first time that $H < h_c$, is also recorded for each ensemble member. Note that $T_h < T_\lambda$ because once the Hamiltonian criterion $H < h_c$ is satisfied (fixing T_h), the vortex must complete its current oscillation before the aspect ratios increases above those allowed in the ACW regime, before eventually reaching $\lambda = \lambda_c$ at T_λ.

Fig. 8 shows a test of the 'rapid fluctuation' results for both the rotation O-U process (top) and the strain O-U process (bottom). For the rotation O-U process, ensembles of 10^4 simulations of the homogenised equations (18), valid in the limit $\delta \to 0$ and with $\kappa = \varepsilon^2\delta = 6.25 \times 10^{-4}$, are compared with the predictions from (33) (solid curves). The solid points show the ensemble mean of T_h in the simulations, with error bars showing 95% confidence intervals. The unfilled points show the mean SSW time T_λ, (in the simulations with $\Gamma_0 = 0.04$ and $\Gamma_0 = 0.06$, $\lambda_c = 4.5$ and 5 respectively). Fig. 8 shows that the theory accurately predicts the mean value of T_h across a wide

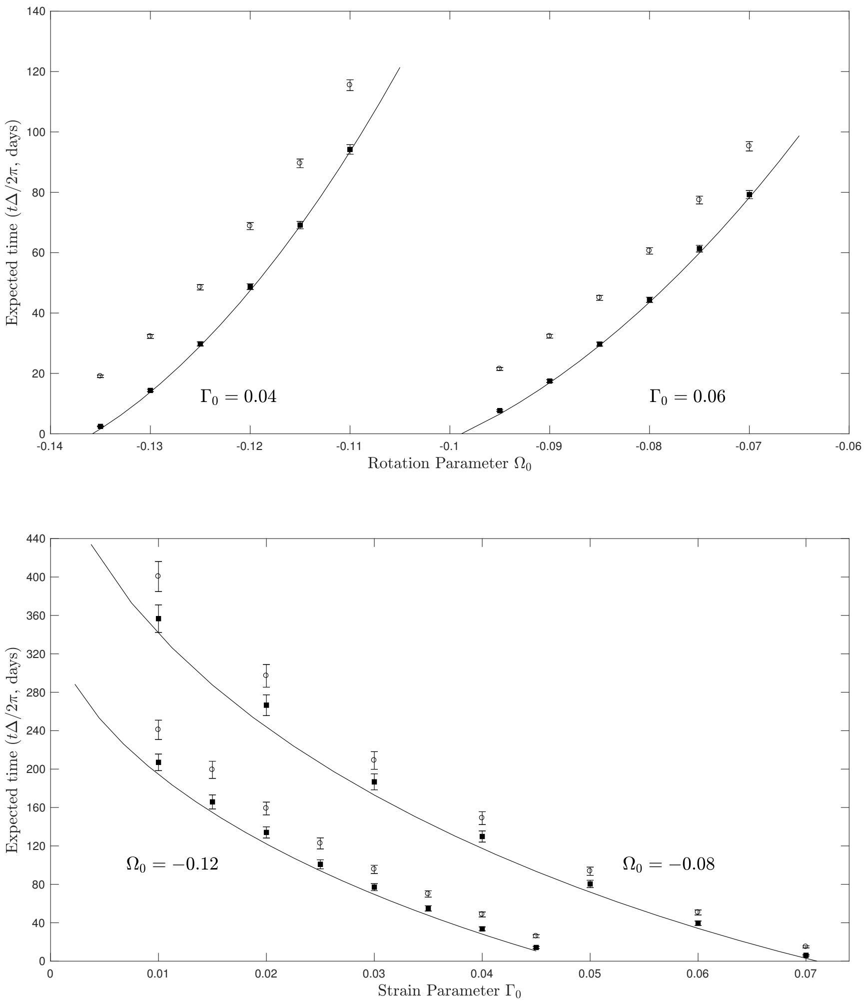

Figure 8. Mean first passage time T_h from the rapid fluctuation theory predictions (solid curves), calculated using (33), and ensemble means of simulations (solid points, error bars show 95% confidence limits), and mean first SSW time T_λ (unfilled points). Top panel. For the rotation O-U process, plotted as a function of the rotation parameter Ω_0 with (left) $\Gamma_0 = 0.04$ and (right) $\Gamma_0 = 0.06$. The results are for the homogenised equations (18), valid for $\delta \to 0$, and with $\kappa = \varepsilon^2\delta = 6.25 \times 10^{-4}$. Bottom panel: For the strain O-U process, as a function the strain parameter Γ_0 with (left) $\Omega_0 = -0.12$ and (right) $\Omega_0 = -0.08$. Here the simulations use Kida's equations coupled to (17), with $\delta = 0.5\pi = \frac{1}{4}$ 'days' and $\varepsilon = 0.0125\delta^{-1}$.

range of parameter values. In these simulations $\Gamma_0 = 0.04$ or 0.06 is fixed, and Ω_0 is varied. The lag between T_h and T_λ of around 20 'days' is approximately constant across the experiments, and is quite a bit longer than the typical oscillation periods in Fig. 4, which reflects the fact that, in the constant parameter situation, the period $T_p \to \infty$ as $H \to h_c$.

The lower panel of Fig. 8 shows mean T_h and T_λ for the strain O-U process near the rapid fluctuation limit. In this case Kida's equations are integrated, along with (17), for parameters $\delta = 0.5\pi = \frac{1}{4}$ 'days' and $\varepsilon = 0.0125\delta^{-1}$, so that $\kappa = 1.5625 \times 10^{-4}$. A smaller ensemble size of 10^3 is used, and in this case Ω_0 is held constant, at either -0.12 or -0.08, while Γ_0 is varied. The smaller ensemble size is necessary as much longer integrations are required when Γ_0 is small. The agreement with the theory for T_h is slightly less good than for the rotation O-U case, due to the finite value of δ, which is nevertheless significantly less than an oscillation period T_p. Comparing the rotation and strain O-U processes at the same parameter setting $(\Omega_0, \Gamma_0) = (-0.12, 0.04)$, the expected time T_λ for an SSW is about the same in each case, despite $\kappa = \varepsilon^2\delta$ being smaller by a factor of 4 in the strain O-U case. In other words, to push the system towards an SSW at the same rate, the noise acting on the strain needs to have only half the amplitude of that acting on the rotation.

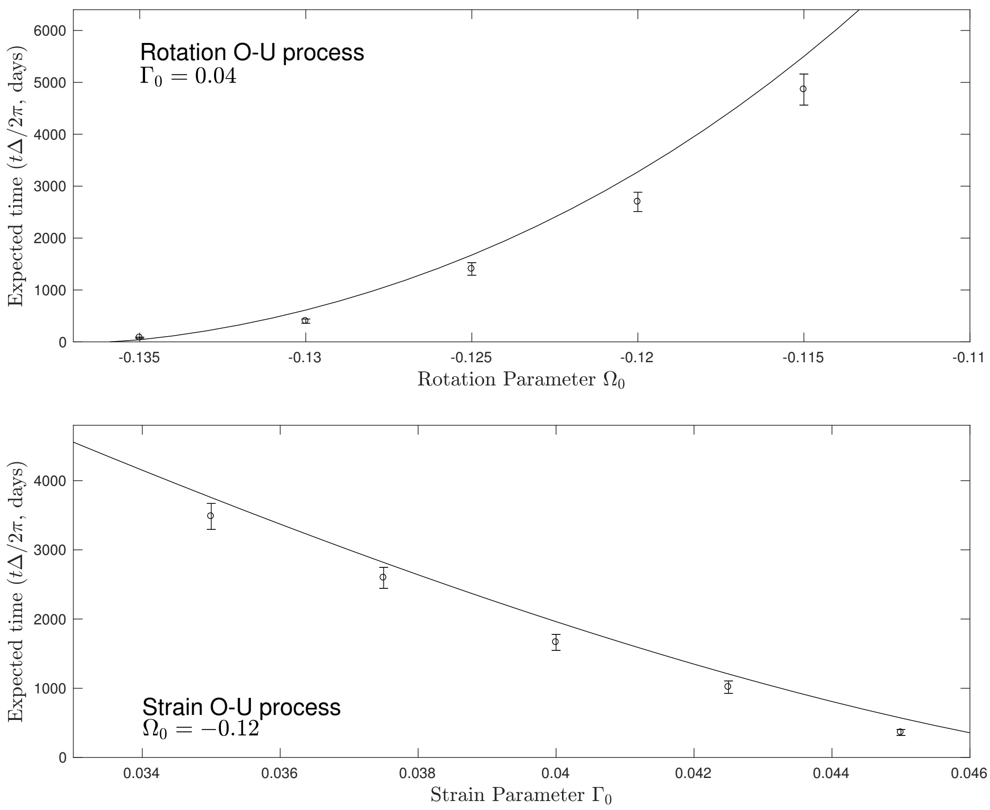

Figure 9. Mean first passage time T_h from the slow evolution theory predictions (solid curves), calculated using (33), and ensemble means of mean first SSW time T_λ (points, error bars show 95% confidence limits) from simulations of Kida's equations. Top panel: For the rotation O-U process, as a function of the rotation parameter Ω_0, with $\Gamma_0 = 0.04$. Bottom panel: For the strain O-U process, as a function of the the strain parameter Γ_0, with $\Omega_0 = -0.12$. In both cases $\delta = 32\pi$ ($= 16$'days'). In in the rotation O-U case $\varepsilon = 0.005$ and in the strain O-U case $\varepsilon = 0.0025$.

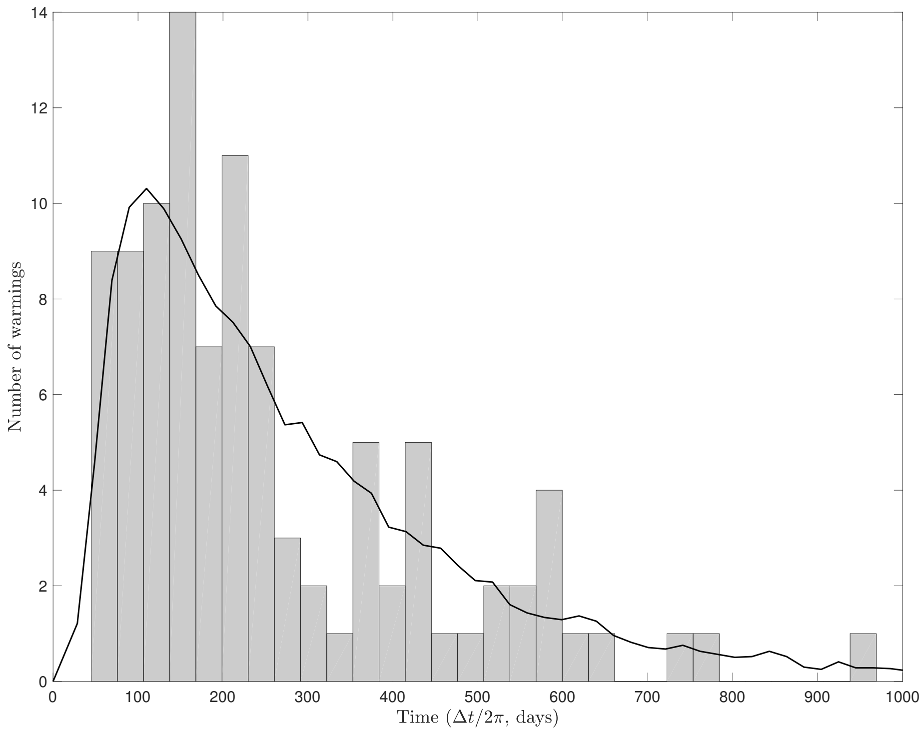

Figure 10. Histogram of SSW onset times T_λ in 100 quasi-geostrophic simulations. The time T_λ is the first time that the vortex aspect ratio $\lambda > 4.5$. The black curve shows the pdf of the SSW time in the corresponding Kida model, calculated using an ensemble of size 10^4.

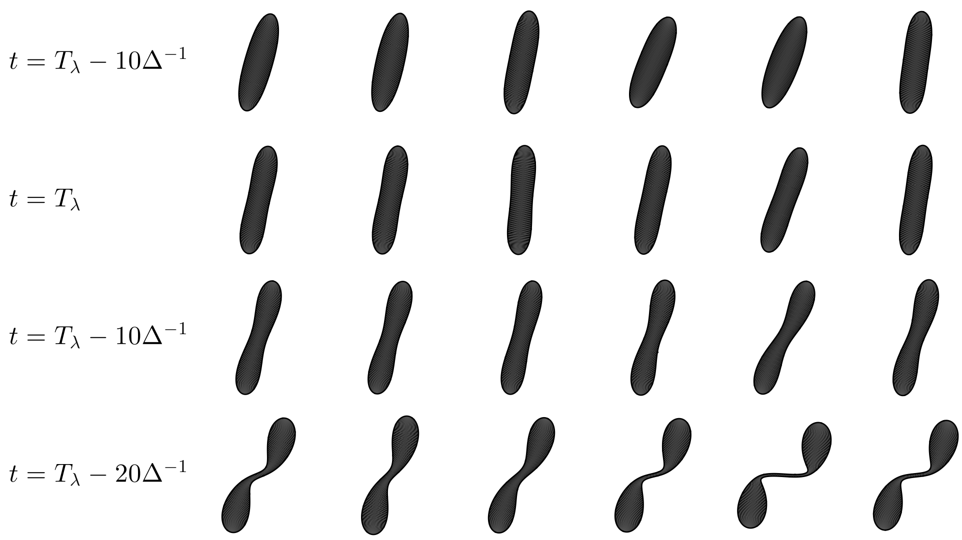

Figure 11. Snapshots of the vortex in 6 quasi-geostrophic simulations, showing vortex splits. The rows show snapshots relative to the SSW time T_λ, defined as the first time that the vortex aspect ratio $\lambda > 4.5$.(First row: $t = T_\lambda - 10\Delta^{-1}$, second: $t = T_\lambda$ third: $t = T_\lambda + 10\Delta^{-1}$ $t = T_\lambda + 20\Delta^{-1}$, or relative model days -1.6,0,+1.6,+3.2 respectively). The times $T_\lambda = \{192, 168, 230, 202, 154, 414\}$ model days respectively.

In Fig. 9, the 'slow evolution' results for the mean first passage time T_λ are tested for both the rotation O-U process (top), plotted as a function of Ω_0 with fixed $\Gamma_0 = 0.04$ and the strain O-U process (bottom), plotted as a function of Γ_0 with fixed $\Omega_0 = -0.12$. In both cases the O-U timescale $\delta = 32\pi$ (16 'days') and $\varepsilon = 0.005$ for the rotation O-U process and $\varepsilon = 0.0025$ for the strain O-U process. At these parameter settings the mean time-scale T_λ for an SSW is rather long, indicating that over a 90 day winter period, SSWs would occur only as a rare event. Consequently only T_λ is plotted, as the relative difference with T_h is small. The basic parameter dependency is well-captured by the theory, which nevertheless seems to overestimate T_λ systematically by 10% or so, which seems to be a finite ε effect.

In summary, the simulations above show that, because the 'climate' parameters (Ω_0, Γ_0) control the critical value h_c that must be attained by the Hamiltonian H in order to trigger an SSW, they can exert significant control over the mean time for an SSW. For example changes in (Ω_0, Γ_0) that act to bring h_c closer to zero (see Fig. 5) have been shown above to reduce the mean time for an SSW significantly (as shown by e.g. Fig. 8).

4.3. *Application to quasi-geostrophic simulations*

To demonstrate the relevance of the results above, the behaviour of a somewhat more realistic quasi-geostrophic model is examined next. The model is the single-layer quasi-geostrophic model of ME11, which solves the quasi-geostrophic potential vorticity equation in an unbounded two-dimensional domain

$$q_t + \mathcal{J}(\psi, q) = 0, \quad q = \nabla^2\psi + h_T, \tag{35}$$

where q is potential vorticity, ψ streamfunction, and h_T a prescribed topography, and the Jacobian operator $\mathcal{J}(f, g) = f_x g_y - f_y g_x$. Exploiting the idea of a 'topographic velocity' discussed in the introduction, an equivalent system that is conceptually closer to the Kida model is

$$q_t + \mathcal{J}(\psi_D + \psi_T, q) = 0, \quad q = \nabla^2\psi_D, \tag{36}$$

where ψ_T is the topographic streamfunction, satisfying $\nabla^2\psi_T = -h_T$, and ψ_D is the dynamic streamfunction determined by q. The conservation properties of (36) are exploited by restricting q to two regions of constant PV, i.e.

$$q(\mathbf{x}, t) = \begin{cases} 1 + 2\Omega(t), & \mathbf{x} \in \mathcal{D} \\ 2\Omega(t), & \mathbf{x} \notin \mathcal{D}. \end{cases} \tag{37}$$

where $\mathcal{D}(t)$ is a time-varying region with constant area (up to numerical error and possible 'contour surgery'), and Ω is the background rotation as in the Kida model. $\mathcal{D}(0)$ is a unit circle centred on the origin. These choices allow (36) to be solved numerically using the contour dynamics with surgery algorithm (Dritschel 1988).

The topography is, in polar coordinates (r, ϕ), given by

$$h_T = h_0(t)\mathrm{J}_2(\gamma r)\cos 2(\phi - \Phi(t)) \tag{38}$$

with the Bessel function form chosen so that the topographic streamfunction is easily obtained as $\psi_T = h_T/\gamma^2$. In the limit of small radial wavenumber $\gamma \to 0$, ψ_T becomes the streamfunction of a strain flow with rate $\Gamma = h_0/4$, and the Kida model is recovered. In order that the model simulates realistic-looking splits, however, we choose finite radial wavenumber $\gamma = 1.162$ (following ME11, with the wavenumber made non-dimensional using the initial unit vortex radius). In this case, the mean strain experienced by the vortex depends weakly on its radius and aspect ratio, with $\Gamma \approx 0.21h_0$ in our model experiments. As the vortex becomes elongated the vortex 'feels' a topographic velocity that deviates significantly from a linear strain flow, and when the bifurcation occurs and the vortex aspect becomes large (i.e. once $\lambda \gtrsim 4.5$), the more complex topographic velocity induces a split. A key difference with ME11, where $\Phi = 0$, is that $\Phi = 2^{1/2}\kappa W$ where $W(t)$ is a Wiener process. The physical interpretation for adding the noise to Φ is not that the physical topography actually rotates, but as a convenient method to access the rapid fluctuation limit ($\delta \to 0$) for the rotation O-U process (16). The rapid fluctuation limit is chosen for investigation because analytical predictions for the mean SSW time can be tested at a relatively cheap computational cost.

An ensemble of 100 simulations, with parameters $h_0 = 0.16$, $\Omega = -0.12$ and $\kappa = 3.125 \times 10^{-4}$ is investigated. Each integration is continued until $T_\lambda + 20\Delta^{-1}$, where T_λ is the first time that $\lambda > 4.5$. These quasi-geostrophic simulations are compared with 10^4 integrations of the stochastic Kida model (19) with $\Gamma = 0.0336 = 0.21h_0$. A histogram of the distribution of SSW times T_λ in the quasi-geostrophic model is shown in Fig. 10, with the solid curve showing the corresponding histogram for the Kida model as a pdf. Good agreement between the models, given the finite quasi-geostrophic ensemble size, is evident. If a winter season is taken to be 100 days (1 day $= 2\pi\Delta^{-1}$), it is notable that each model is in a reasonably realistic regime in the sense that the probability of an SSW occurring within the season is around 18%.

Fig. 11 shows snapshots of the vortex at times close to T_λ for the first six simulations. Following the interpretation of section 3.4.1 the vortex is plotted relative to the topography. The snapshots show that:

1. Despite T_λ being realised at widely varying times (between 154 and 414 model days) a similar-looking vortex split invariably follows.
2. The time taken for the split to develop following T_λ is short (the final row shows $T_\lambda + 3.2$ days), although stochasticity introduces noticeable variation between the simulations in the time taken for a split to occur.
3. The orientation of the vortex elongation and subsequent split, measured relative to the underlying topography, remains remarkably similar between simulations.

Observed vortex split SSWs in the Northern hemisphere share each of the features 1-3 described above (Matthewman *et al.* 2009).

5. Conclusions

The main contribution of this work has been to introduce a simple model which demonstrates that vortex splitting SSWs can result from the cumulative effects of weak 'noise'. In the simple model, the SSWs occur because the noise induces a random walk (with drift) in the vortex Hamiltonian H, and this random walk can cause H to reach a critical value h_c, which corresponds to a bifurcation in the periodic orbits of the model. The noise in question can be identified with unsteadiness in tropospheric planetary wave forcing, i.e. tropospheric macroturbulence. Extrapolating this picture, Antarctic winters featuring large oscillations in vortex aspect ratio (e.g. 2012, 2013, 2016) correspond to realisations in which H becomes negative, and winters without significant oscillations (e.g. 2014 and 2015) have H positive. Further, the SSW of 2002 is a rare event in which $H < h_c$, the (negative) critical value associated with an SSW in Kida's model. The oscillations in aspect ratio appear to be essentially the stratospheric vacillations discovered by Scaife *et al.* (2005) which, interestingly, appear to have a strongly nonlinear vortex-splitting counterpart (Scott 2016).

Mathematical analysis of the simple model reveals the following:

1. When the noise takes the form of an O-U process driving the linear flow in Kida's model, the random walk with drift in H can be derived analytically in two distinct limits. The first passage time problem for $H < h_c$ can then be solved, with the expected time T_λ for an SSW found from the result.
2. The expected time T_λ for an SSW can be found as a function of the parameters describing the background flow and O-U process. The timescale T_λ depends strongly on the critical value h_c for the bifurcation. Broadly speaking, the further h_c is away from zero, the longer it will take the random walk to reach it.
3. In terms of causing an SSW, an O-U process forcing the strain component of the background flow is over twice as efficient compared to one forcing the rotation component, in the sense that T_λ is smaller in the former case at even at half the forcing amplitude of the latter.
4. Numerical simulations show an O-U process, at fixed amplitude ε, is most efficient at causing an SSW when the decorrelation timescale $\delta \sim 0.1 - 0.2T_p$ where T_p is the oscillation period.

Overall, the results point towards a 'noise-memory' paradigm for the winter stratosphere, in which the current state of the stratosphere, represented in the simple model by the Hamiltonian H, depends on the history of the forcing over a significant period. Even in the simple model, the precise dependence on the forcing history is opaque, and in particular it is to be emphasised that large forcing amplitudes are not necessary to bring about an SSW. Attempts to search for the dynamical 'cause' of an SSW, for example by analysing Rossby wave activity in the troposphere in the lead-up, may therefore be unproductive. Many previous authors have discussed 'pre-conditioning' of the vortex before an SSW. The noise-memory paradigm supports the idea of pre-conditioning, but suggests that what is important is changes to the dynamical state of the vortex (as measured in our model by H), as opposed to its changes in its physical structure.

Extrapolating the results of our simple model to the stratospheric vortices, SSW frequency is particularly sensitive to climatic changes which act to reduce the background zonal wind at the vortex edge (i.e. lower Ω, see Fig. 8) as the vortex will be brought closer to nonlinear resonance. Climatic changes that act to increase *fluctuations* in forcing, e.g. due to more active tropospheric storm tracks, are also particularly effective at increasing SSW frequency (e.g. Fig. 6). A major caveat is that physics missing from the simple model must naturally also be considered. For example there is no representation of the season cycle, radiative damping, or momentum fluxes from gravity waves, for which there is increasing evidence of an important role (e.g. Albers and Birner 2014) in individual SSW events. Further modelling studies are required to investigate the importance each of these effects, although speculatively it seems likely that the various forcings will act mainly to determine the (time-dependent) parameter regime for the polar vortices, and no doubt to limit the time-scale over which the noise-memory persists. To gain a more quantitative description it will be also necessary to re-introduce vertical structure and more realistic topographic forcing into the model. In the Arctic in particular, it is unlikely that the simple model offers more than qualitative insight, as the changing vertical structure of the Arctic vortex, as well as large horizontal migrations of the vortex centroid have too strong an influence on the dynamics. In the Antarctic, however, there is tentative evidence that Kida's model may have useful predictive power. The question of how best to 'fit' the parameters of Kida's model, and other models in the model hierarchy of SSWs, to the observations will be the subject of future work.

Acknowledgements

JGE acknowledges support from the Leverhulme trust RF-2016-158. MM is funded by an EPSRC studentship.

A. Derivation of the cycle-averaged equation

For simplicity, we present the derivation of the cycle-averaged equation (14) for a slightly simplified example in which the only non-zero component of stochastic forcing (9) is on the rotation variable Ω, i.e.

$$\mathrm{d}\Omega = \varepsilon F^{\Omega}(\Omega)\mathrm{d}t + \varepsilon^{1/2}\Sigma^{\Omega}(\Omega)\mathrm{d}W_3. \tag{39}$$

The results for the more general forcing case (9) follow by exact analogy.

The FPE for the system (39) coupled with (10-11) describes the time-evolution of the probability density $p(\lambda, h, \omega, t)$ associated with the random variables $\{\Lambda, H, \Omega\}$. Following standard techniques (e.g. §3.4.1 of Gardiner 2009), the FPE is

$$\begin{aligned} p_t + \left((-V(\lambda,h,\omega))^{1/2}p\right)_\lambda - \varepsilon\left(F^{\Omega}(\omega)\frac{(\lambda-1)^2}{\lambda}p\right)_h \\ + \varepsilon\left(F^{\Omega}(\omega)p\right)_\omega = \frac{\varepsilon}{2}\Bigg(\left(\Sigma^{\Omega}(\omega)^2 p\right)_{\omega\omega} \\ -2\left(\Sigma^{\Omega}(\omega)^2\frac{(\lambda-1)^2}{\lambda}p\right)_{\omega h} + \left(\Sigma^{\Omega}(\omega)^2\frac{(\lambda-1)^4}{\lambda^2}p\right)_{hh}\Bigg), \end{aligned} \tag{40}$$

where subscripts denote partial derivates. The correct interpretation of the square root in (40) is that the λ-domain for p, $\lambda \in [\lambda_-(h,\omega), \lambda_+(h,\omega)]$ is in fact doubled, with one part-solution p^+ taking the positive branch of the square root and the other p^- the negative branch. The two parts of the solution, which are associated with the increasing and decreasing phases of the vortex oscillation respectively, communicate through the probability flux conditions $(-V(\lambda_\pm))^{1/2}p^+(\lambda_\pm) = (-V(\lambda_\pm))^{1/2}p^-(\lambda_\pm)$ at $\lambda = \lambda_\pm$.

The method of multiple-scales can be applied to (40) by seeking a solution based on an ansatz of the form

$$p = p_0(\lambda, h, \omega, t, \tau) + \varepsilon p_1(\lambda, h, \omega, t, \tau) + \ldots \tag{41}$$

where $\tau = \varepsilon t$ is a 'slow' time-scale associated with many periods of the vortex oscillation. Introduction of the slow time-scale requires

$$\frac{\partial}{\partial t} \to \frac{\partial}{\partial t} + \varepsilon \frac{\partial}{\partial \tau}. \tag{42}$$

Inserting the ansatz (41) into (40) gives at leading order

$$p_{0t} + \left((-V(\lambda, h, \omega))^{1/2} p_0 \right)_\lambda = 0. \tag{43}$$

This equation can be solved for p_0 by transforming variables (λ, t) to characteristic variables $(\tilde{\lambda}, \eta)$, where

$$\tilde{\lambda} = \lambda, \quad \eta = t - T_k^{\pm}(\lambda, h, \omega), \tag{44}$$

where $T_k^{\pm}$ is the multi-valued oscillation time

$$T_k^{\pm}(\lambda) = \int_{C_k^{\pm}(\lambda)} \frac{\mathrm{d}q}{(-V(q))^{1/2}}, \tag{45}$$

In this definition the possible integration paths $C_k^{\pm}(\lambda)$ follow C, starting at λ_- and finish at λ, with the positive sign corresponding to arriving at λ on the upper branch, the negative sign the lower branch, and $k \geq 0$ denoting the number of completed oscillations. Evidently, because of the periodicity of the oscillation, $T_k^{\pm}(\lambda) = T_0^{\pm}(\lambda) + kT_p$. The fact that the function $T_k^{\pm}$ is multi-valued means that the two branches of the solution of (43) are unfolded by this change of variables, and also shows that the resulting solution is T_p-periodic in η. The general solution for p_0 is then (dropping the tilde on λ)

$$p_0 = \frac{\tilde{P}(\eta, h, \omega, \tau)}{(-V(\lambda, h, \omega))^{1/2}}, \tag{46}$$

for an arbitrary function $\tilde{P}$.

The next order in the expansion of (40) gives

$$\begin{aligned} p_{1t} + \left((-V(\lambda, h, \omega))^{1/2} p_1 \right)_\lambda = \\ - p_{0\tau} + \left(F^\Omega(\omega) \frac{(\lambda-1)^2}{\lambda} p_0 \right)_h - \left(F^\Omega(\omega) p_0 \right)_\omega \\ + \frac{1}{2} \Bigg(\left(\Sigma^\Omega(\omega)^2 p_0 \right)_{\omega\omega} - 2 \left(\Sigma^\Omega(\omega)^2 \frac{(\lambda-1)^2}{\lambda} p_0 \right)_{\omega h} \\ + \left(\Sigma^\Omega(\omega)^2 \frac{(\lambda-1)^4}{\lambda^2} p_0 \right)_{hh} \Bigg). \end{aligned} \tag{47}$$

To obtain an equation for the long-time evolution of the system it is not necessary to solve for p_1. Instead, it is sufficient to apply both a time (t)-average, and cycle integral $\oint_C \cdot \, \mathrm{d}\lambda$ to (47), which remove the terms involving p_1. Denoting the time-average of $\tilde{P}$ by

$$P(h, \omega, \tau) = \lim_{t_m \to \infty} \frac{1}{t_m} \int_0^{t_m} \tilde{P} \, \mathrm{d}t = \frac{1}{T_p} \int_0^{T_p} \tilde{P}(\eta) \, \mathrm{d}\eta, \tag{48}$$

the averaging results in the following 'slow-evolution' equation for P,

$$\begin{aligned} & P_\tau - \left(F^\Omega \langle G^\Omega \rangle P \right)_h + \left(F^\Omega P \right)_\omega \\ & = \tfrac{1}{2} \left(\Sigma^{\Omega 2} P \right)_{\omega\omega} - \left(\Sigma^{\Omega 2} \langle G^\Omega \rangle P \right)_{\omega h} + \tfrac{1}{2} \left(\Sigma^{\Omega 2} \langle G^{\Omega 2} \rangle P \right)_{hh}. \end{aligned} \tag{49}$$

where $G^\Omega = (\lambda - 1)^2/\lambda$, and $\langle \cdot \rangle$ denotes the cycle average.

Equation (49) is the FPE of the following coupled stochastic process in (H, Ω)

$$\mathrm{d}\Omega = F^\Omega(\Omega) \mathrm{d}\tau + \Sigma^\Omega(\Omega) \mathrm{d}B_3, \tag{50}$$

$$\begin{aligned} \mathrm{d}H = & -F^\Omega(\Omega) \langle G^\Omega \rangle (H, \Omega) \mathrm{d}\tau - \Sigma^\Omega(\Omega) \langle G^\Omega \rangle (H, \Omega) \mathrm{d}B_3 \\ & + \Sigma^\Omega(\Omega) \langle \langle G^\Omega \rangle \rangle (H, \Omega)^{1/2} \mathrm{d}B. \end{aligned} \tag{51}$$

where B_3 and B are independent Wiener processes in the slow time variable τ. Applying the methodology above using the more general forcing (9), leads directly to the cycle-averaged equation (14).

B. Homogenisation applied to O-U forcing in Kida's equations

In this section the mathematical method for homogenisation of O-U processes is presented, following e.g. the treatment in Pavliotis and Stuart (2007). Two examples are covered in detail.

B.1. Homogenisation of Kida's equations

Consider first Kida's equations (1) coupled to the O-U process (16) for Ω. Introducing $\bar{\Omega} = (\Omega - \Omega_0)/\varepsilon$, and substituting $\varepsilon^2\delta = \kappa$ gives (taking $\Phi = 0$ without loss of generality)

$$\mathrm{d}\bar{\Omega} = -\delta^{-1}\bar{\Omega}\mathrm{d}t + 2^{1/2}\delta^{-1/2}\mathrm{d}W_3$$

$$\mathrm{d}\Theta = \Bigg(\Omega_0 + \delta^{-1/2}\kappa^{1/2}\bar{\Omega} \tag{52}$$

$$+ \frac{\Lambda}{(\Lambda+1)^2} - \frac{\Lambda^2+1}{\Lambda^2-1}\Gamma\sin 2\Theta\Bigg)\,\mathrm{d}t \tag{53}$$

$$\mathrm{d}\Lambda = 2\Lambda\Gamma\cos 2\Theta\,\mathrm{d}t.$$

The FPE describing the time-evolution of the probability density $p(\lambda,\theta,\omega,t)$ of the random variables $\{\Lambda,\Theta,\bar{\Omega}\}$ is therefore

$$p_t - \delta^{-1}(\omega p)_\omega + \delta^{-1/2}\kappa_3^{1/2}(\omega p)_\theta + (f(\lambda,\theta)p)_\theta + (g(\lambda,\theta)p)_\lambda = \delta^{-1}p_{\omega\omega}, \tag{54}$$

where

$$f(\lambda,\theta) = \Omega_0 + \frac{\lambda}{(\lambda+1)^2} - \frac{\lambda^2+1}{\lambda^2-1}\Gamma\sin 2\theta, \tag{55}$$

$$g(\lambda,\theta) = 2\lambda\Gamma\cos 2\theta. \tag{56}$$

Homogenisation theory describes the asymptotic behaviour of (54) when $\delta \to 0$. To proceed, a solution of (54) is sought as a power series in $\delta^{1/2}$,

$$p = p_0(\lambda,\theta,\omega,t) + \delta^{1/2}p_1(\lambda,\theta,\omega,t) + \delta p_2(\lambda,\theta,\omega,t) + \ldots, \tag{57}$$

At leading order $\mathcal{L}p_0 = 0$, where the linear operator $\mathcal{L}$ acts on functions $h(\omega)$ according to $\mathcal{L}h = h_{\omega\omega} + (\omega h)_\omega$. The general solution, using the condition that p is integrable in ω, is

$$p_0 = P(\lambda,\theta,t)\mathrm{e}^{-\omega^2/2}. \tag{58}$$

At the next order, the equation is

$$\mathcal{L}p_1 = \kappa^{1/2}(\omega p_0)_\theta\,, \tag{59}$$

which has solution

$$p_1 = -\kappa^{1/2}P_\theta(\lambda,\theta,t)\omega\mathrm{e}^{-\omega^2/2}. \tag{60}$$

To complete the theory, the next order equation must also be considered,

$$\mathcal{L}p_2 = p_{0t} - (\omega p_0)_\omega + \kappa^{1/2}(\omega p_1)_\theta + (f(\lambda,\theta)p_0)_\theta + (g(\lambda,\theta)p_0)_\lambda\,. \tag{61}$$

It is not necessary to solve explicitly for p_2. Instead, the solvability condition of (61) can be used to obtain an equation for P. The solvability condition is applied by substituting for p_0 and p_1 and integrating (61) in ω. The result is

$$P_t + (f(\lambda,\theta)P)_\theta + (g(\lambda,\theta)P)_\lambda = \kappa P_{\theta\theta}. \tag{62}$$

Equation (62) is the FPE of the homogenised system (18).

B.2. Homogenisation of the cycle-averaged equations

Next, homogenisation is used to obtain the long-time behaviour of the cycle-averaged equation (21). To proceed we need to exploit the fact that $\varepsilon \ll 1$ and define $\bar{\Omega} = (\Omega - \Omega_0)/\varepsilon$. The FPE for the pdf $p(\omega,h,\tau)$ of $\{\bar{\Omega},H\}$ is, to $O(\varepsilon^2)$ in accuracy

$$p_\tau - (\omega p)_\omega + \varepsilon\left(\omega\langle G^\Omega\rangle_0 p\right)_h + \varepsilon^2\left(\omega^2\partial_\omega\langle G^\Omega\rangle|_0 p\right)_h = p_{\omega\omega} - 2\varepsilon\left(\langle G^\Omega\rangle_0 p\right)_{\omega h} + \varepsilon^2\left(\langle G^{\Omega^2}\rangle_0 p\right)_{hh} \quad (63)$$

Seeking a solution

$$p = p_0(\omega, h, \bar\tau) + \bar\varepsilon p_1(\omega, h, \bar\tau) + ... \quad (64)$$

where $\bar\tau = \bar\varepsilon^2\tau$ is a long time-scale, gives $\mathcal{L}p_0 = 0$ at leading order and $p_0 = P(h, \bar\tau)\mathrm{e}^{-\omega^2/2}$. At first order

$$\mathcal{L}p_1 = 2\left(G_0^\Omega(h)p_0\right)_{\omega h} + \left(\omega G_0^\Omega(h)p_0\right)_h, \quad (65)$$

which has solution

$$p_1 = \left(G_0^\Omega(h)P\right)_h \omega\mathrm{e}^{-\omega^2/2}. \quad (66)$$

At second order,

$$\mathcal{L}p_2 = 2\left(\langle G^\Omega\rangle_0 p_1\right)_{\omega h} + \left(\omega\langle G^\Omega\rangle_0 p_1\right)_h + p_{0\bar\tau} + \left(\omega^2\partial_\omega\langle G^\Omega\rangle|_0 p_0\right)_h - \left(\langle G^{\Omega^2}\rangle_0 p_0\right)_{hh}. \quad (67)$$

Inserting for p_0 and p_1, and applying the solvability condition by integrating in ω, gives

$$P_{\bar\tau} + \left(\langle G^\Omega\rangle_0\partial_h\left(\langle G^\Omega\rangle_0 P\right)_h\right)_h + \left(\partial_\omega\langle G^\Omega\rangle|_0 P\right)_h - \left(\langle G^{\Omega^2}\rangle_0 P\right)_{hh} = 0, \quad (68)$$

which can be seen, after substituting for $\bar\tau$ and some rearrangement, to be the FPE of (22).

C. Details of the first passage time problem

Here the details of the first passage time problem for equation (30) are presented (following e.g. section 5.2.7 of Gardiner 2009). First, it is useful to define $p(h, t, h', t')$ to be the probability density of $H(t) \in (h_c, h_m)$, given the deterministic initial condition $H(t') = h'$. An 'absorbing' boundary for (30) is applied at $H = h_c$, and a reflecting boundary at $H = h_m$, because we are interested in finding the expected time at which H is absorbed at the boundary $H = h_c$.

In addition to the 'forwards' FPE, p satisfies the backwards Kolmogorov equation (BKE)

$$p_{t'} = -a(h')p_{h'} - \tfrac{1}{2}b(h')p_{h'h'}. \quad (69)$$

To determine the first passage time, it is helpful to consider $G(h', t') = p(h_c, t', h', 0) = p(h_c, 0, h', -t')$, which is the probability density of first reaching h_c at time t', starting at $H(0) = h'$. The second expression for $G(h', t')$ follows from the fact that the process (30) is stationary, and that consequently p can only depend on its time arguments in the combination $t - t'$. The expression for G can be inserted into the BKE to give

$$G_{t'} = a(h')G_{h'} + \tfrac{1}{2}b(h')G_{h'h'}, \quad (70)$$

with the associated boundary conditions

$$G(h_c, t') = 0, \quad G_{h'}(h_m, t') = 0. \quad (71)$$

The boundary conditions correspond to absorption at $h' = h_c$ and reflection at $h' = h_m$, since H is confined to the domain $h_c < H < h_m$, but can only 'escape' at h_c.

The first passage time of interest can now be defined as the expectation

$$T(h') = \int_0^\infty t' G(h', t')\,\mathrm{d}t'. \quad (72)$$

Multiplying (70) by t', integrating, and using the fact that

$$\int_0^\infty G(h', t')\,\mathrm{d}t' = 1, \quad (73)$$

leads directly to equation (31) (in which primes have been dropped).

References

Albers JR, Birner T. 2014. Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings. *J. Atmos. Sci.* **71**(11): 4028–4054.

Ayarzagüena B, Langematz U, Meul S, Oberländer S, Abalichin J, Kubin A. 2013. The role of climate change and ozone recovery for the future timing of major stratospheric warmings. *Geophysical Research Letters* **40**(10): 2460–2465.

Baldwin MP, Dunkerton TJ. 2001. Stratospheric harbingers of anomalous weather regimes. *Science* **294**: 581–584.

Birner T, Williams PD. 2008. Sudden stratospheric warmings as noise-induced transitions. *J. Atmos. Sci.* **65**: 3337–3343.

Butchart N, Austin J, Knight JR, Scaife AA, Gallani ML. 2000. The Response of the Stratospheric Climate to Projected Changes in the Concentrations of Well-Mixed Greenhouse Gases from 1992 to 2051 . *J. Clim.* **13**: 2142–2159.

Charlton AJ, Polvani LM. 2007. A new look at stratospheric sudden warmings I: climatology and modeling benchmarks. *J. Clim.* **20**: 449–469.

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hlm EV, Isaksen L, Kallberg P, Kohler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut J, Vitart F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**(656): 553–597.

Dritschel DG. 1988. Contour surgery: a topological reconnection scheme for extended integrations using contour dynamics. *J. Comput. Phys.* **77**: 240–266.

Dritschel DG. 1990. The stability of elliptical vortices in an external straining flow. *J. Fluid Mech.* **210**: 223–261.

Esler JG, Matthewman NJ. 2011. Stratospheric sudden warmings as self-tuning resonances. Part II: vortex displacement events. *J. Atmos. Sci.* **68**: 2505–2523.

Esler JG, Polvani LM, Scott RK. 2006. The antarctic sudden stratospheric warming of 2002: A self-tuned resonance? *Geophys. Res. Lett.* **33**. Doi:10.1029/2006GL026034.

Esler JG, Scott RK. 2005. Excitation of transient Rossby waves on the stratospheric polar vortex and the barotropic sudden warming. *J. Atmos. Sci.* **62**: 3661–3682.

Gardiner CW. 2009. *Stochastic methods: A handbook for the natural and social sciences*. Springer: Berlin, 4 edn. 447pp.

Held IM. 1999. The macroturbulence of the troposphere. *Tellus A* **51**: 59–70.

Holton JR, Mass C. 1976. Stratospheric vacillation cycles. *J. Atmos. Sci.* **33**: 2218–2225.

Kida S. 1981. Motion of an Elliptic Vortex in a Uniform Shear Flow. *J. Phys. Soc. Jap.* **50**: 3517–3520.

Liu YS, Scott RK. 2015. The onset of the barotropic sudden warming in a global model. *Q. J. R. Meteorol. Soc.* **141**: 2944?2955.

Matthewman NJ, Esler JG. 2011. Stratospheric sudden warmings as self-tuning resonances. Part I: vortex splitting events. *J. Atmos. Sci.* **68**: 2481–2504.

Matthewman NJ, Esler JG, Charlton-Perez AJ, Polvani LM. 2009. A new look at stratospheric sudden warmings. Part III. Polar vortex evolution and vertical structure. *J. Clim.* **22**: 1566–1585.

Maycock AC, Hitchcock P. 2015. Do split and displacement sudden stratospheric warmings have different annular mode signatures? *Geophys. Res. Lett.* **42**(24): 10,943–10,951. 2015GL066754.

Mitchell DM, Charlton-Perez AJ, Gray LJ. 2011. Characterizing the variability and extremes of the stratospheric polar vortices using 2D moment analysis. *J. Atmos. Sci.* **68**: 1194–1213.

Mitchell DM, Charlton-Perez AJ, Gray LJ, Akiyoshi H, Butchart N, Hardiman SC, Morgenstern O, Nakamura T, Rozanov E, Shibata K, Smale D, Yamashita Y. 2012a. The nature of Arctic polar vortices in chemistry?climate models. *Q. J. R. Meteorol. Soc.* **138**: 1681–1691.

Mitchell DM, Gray LJ, Anstey J, M P Baldwin AJCP. 2013. The influence of stratospheric vortex displacements and splits on surface climate. *J. Climate* **26**: 2668–2682.

Mitchell DM, Osprey S, Gray LJ, Butchart N, Hardiman S, Charlton-Perez AJ, Watson P. 2012b. The effect of climate change on the variability of the northern hemisphere stratospheric polar vortex. *J. Atmos. Sci.* **69**: 2608–2618.

Nakagawa K, Yamazaki K. 2006. What kind of stratospheric sudden warming propagates to the troposphere? *Geophys. Res. Lett.* **33**: L04 801.

Nielsen-Gammon JW, Lefevre RJ. 1996. Piecewise tendency diagnosis of dynamical processes governing the development of an upper-tropospheric mobile trough. *J. Atmos. Sci.* **53**: 3120–3142.

Pavliotis GA, Stuart AM. 2007. *Multiscale methods averaging and homogenization*. Springer: New York.

Plumb RA. 1981. Instability of the distorted polar night vortex: A theory of stratospheric warmings. *J. Atmos. Sci.* **38**: 2514–2531.

Ruzmaikan A, Lawrence J, Cadavid C. 2003. A simple model of stratospheric dynamics including solar variability. *J. Climate* **16**: 1593?1600,.

Scaife AA, Jackson DR, Swinbank R, Butchart N, Thornton HE, Keil M, Henderson L. 2005. Stratospheric Vacillations and the Major Warming over Antarctica in 2002 . *J. Atmos. Sci.* **62**: 629–639.

Scinocca JF, Haynes PH. 1998. Dynamical forcing of stratospheric waves by the tropospheric circulation. *J. Atmos. Sci.* **55**: 2361–2392.

Scott RK. 2016. A new class of vacillations of the stratospheric polar vortex. *Q. J. R. Meteorol. Soc.* **142**(698): 1948–1957.

Waugh DW. 1997. Elliptical diagnostics of stratospheric polar vortices. *Q. J. R. Meteorol. Soc.* **123**: 1725–1748.

Waugh DW, Randel WJ. 1999. Climatology of Arctic and Antarctic polar vortices using elliptic diagnostics. *J. Atmos. Sci.* **56**: 1594–1613.

