Statistical Physics Interpretation of Southern Ocean Mesoscale Turbulence



Tomos W. David

Jesus College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Atmospheric, Ocean and Planetary Physics

Trinity 2017

Acknowledgements

I begin by thanking Laure Zanna and David Marshall for their considerable support, mentoring, and supervision; it has been an immense privilege and joy to work with them and to learn from them over these past years.

The Ocean Climate group has been great fun to be part of and I would like to thank all its members, past and present, for the many enlightening conversations which has formed a large part of my continuing educating.

In particular, Rob Fraser, my long suffering office mate, has been a constant companion throughout the highs and lows of our DPhil endeavors, both in the office and in more relaxed environments. It is with some sadness, at least for me, I consider our office sharing coming to an end.

With respect to the particulars of this thesis I wish to thank the following people for their greatly valued assistance: Louis-Philippe Nadeau and Jeff Blundell for their help in my journey to understanding momentum conserving boundary conditions for quasi-geostrophic dynamics in multi-connected domains; Geoff Stanley, another comrade and fellow DPhil student, for his generous assistance with obtaining and analyzing the Southern Ocean State Estimate data; Antoine Venaille and Freddy Bouchet for insightful conversations on the statistical mechanics of geophysical flows; and Chris O'Reilly for teaching me the significance test used in this thesis.

This DPhil project was funded by the Natural Environment Research Council and I am also grateful for the travel support provided by Jesus College, where I have spent eight enjoyable years.

Being responsible for both seeding my interest in the ocean as an application of physics as well as my next steps as a researcher I owe Mattias Green immense thanks. The summer work I did with Mattias as an undergraduate was fun, educational and led to places I had not considered at the time.

I would like to extend my great appreciation to both of my examiners, Valerio Lucarini and Andrew Wells, for agreeing to examine this thesis, without whom this process would not be possible.

I will be forever indebted to my father, Huw, mother, Nerys, and brother, Tudur, for their continued support and encouragement. Without the education my parents provided I would not be where I am now.

Finally, and certainly not least, I would like to thank my fiancée, Chelsey, for her unwavering support, encouragement, compassion, enthusiasm, inspiration and understanding throughout this time and for many years before. I simply could not have done this without her.

Abstract

Ocean mesoscale (10-100km) turbulence is integral to the complex and interconnected climate system. Many of the assumptions made in classical turbulence theory are inappropriate for describing the characteristic features of geophysical flows. In particular, when the fluid flow is both forced and dissipative. With increased interest in stochastic parameterization of ocean eddies, it becomes pertinent to consider what the statistics of ocean turbulence should look like as well as consider what statistical physics concepts can be used to further the theoretical understanding of the ocean mesoscale. In this thesis, we examine mesoscale turbulence from both a theoretical and empirical perspective employing both idealized models and realistic ocean model data. The statistical nature of a turbulent barotropic jet is explored empirically in the context of an idealized model as well as for realistic ocean model data. Additionally, an entropic and statistical mechanics framework is investigated in the context of force-dissipative flow. It is shown that: ocean jets which exhibit strong mixing barriers have a profound effect on the statistics, leading to multi-modal probability distributions and pointing to the importance asymmetry and extreme values in the statistics; an eddy-mixing entropy, which quantifies turbulent disorder, is a useful metric in describing turbulent motions and evidence is found for the maximization of this entropy in a forced-dissipative fluid flow. Further, across the separate studies presented in this thesis, it is shown that Lagrangian information is useful in distinguishing large-scale coherent structures from the more stochastic small-scale turbulence; and high-order Casimir invariants of motion are important for describing forced-dissipative geophysical flows in a statistical mechanics framework.

Contents

1 Introduction **1**
- 1.1 Motivation . . . 1
- 1.2 Ocean mesoscale turbulence . . . 3
  - 1.2.1 The Southern Ocean and eddy saturation . . . 4
  - 1.2.2 Stochastic Parameterization . . . 6
- 1.3 Approach of this thesis . . . 7

2 Geophysical fluid dynamics background **9**
- 2.1 Introduction . . . 9
- 2.2 Potential vorticity . . . 10
- 2.3 Quasi-geostrophic dynamics . . . 11
- 2.4 Invariants of motion . . . 15
- 2.5 Hamiltonian formalism . . . 16
- 2.6 The barotropic channel model . . . 18
- 2.7 Dynamics of jets, potential vorticity steps and mixing barriers . . . 20

3 Statistical theories of geostrophic turbulence **22**
- 3.1 Introduction . . . 22
- 3.2 The homogeneous turbulence paradigm . . . 23
  - 3.2.1 Three-dimensional turbulence . . . 23
  - 3.2.2 Two-dimensional turbulence . . . 26
  - 3.2.3 Application to the geophysical flows . . . 28
- 3.3 Miller-Robert-Sommeria statistical mechanics . . . 29
  - 3.3.1 Statistical mechanics and turbulence . . . 29
  - 3.3.2 Eddy-mixing entropy . . . 31
  - 3.3.3 Maximization of entropy . . . 33
- 3.4 Thermodynamic versus information entropy . . . 35
- 3.5 Other applications of entropy and statistical mechanics to climate problems . . . 36

4 The statistical nature of barotropic ocean jets **39**
- 4.1 Introduction . . . 39
- 4.2 Numerical experiments . . . 42
- 4.3 Linear stability and west/east asymmetry . . . 42
- 4.4 Scaling of the mean and standard deviation . . . 48

4.5 Skewness, kurtosis and jet meandering . . . 54
4.6 Streamfunction-following statistics . . . 60
4.7 Discussion . . . 65
4.7.1 Joint probability distributions . . . 65
4.7.2 Mixed distributions . . . 67
4.7.3 Stochastic parameterization . . . 69
4.8 Concluding remarks . . . 70

5 Statistical perspectives on the Southern Ocean 74
5.1 Introduction . . . 74
5.1.1 Potential vorticity front identification . . . 76
5.1.2 Mixing suppression by mean flows . . . 77
5.2 The Southern Ocean State Estimate . . . 77
5.3 Simple potential vorticity statistics: skewness and kurtosis . . . 80
5.3.1 Surface statistics . . . 80
5.3.2 Eulerian statistics . . . 82
5.3.3 Neutral density statistics . . . 85
5.4 Osborn-Cox diffusivity . . . 87
5.5 Distributions of potential vorticity . . . 92
5.5.1 Global distributions of potential vorticity on neutral density surfaces . . . 92
5.5.2 Geographic variation in potential vorticity distributions . . . 97
5.6 Concluding remarks . . . 99

6 Eddy-mixing entropy as a measure of turbulent disorder in barotropic ocean jets 102
6.1 Introduction . . . 102
6.2 Numerical experiments . . . 103
6.3 Analytical model for evolution for entropy . . . 105
6.4 Entropy in the numerical simulations . . . 107
6.4.1 Eddy-mixing entropy in freely-decaying turbulence . . . 107
6.4.2 Eddy-mixing entropy in forced-dissipative turbulence . . . 109
6.5 Relation to the maximum entropy principle . . . 113
6.5.1 Time evolution of entropy and the maximum entropy principle 113
6.5.2 Reconstruction of entropy evolution . . . 115
6.5.3 Solving for the Lagrange multipliers . . . 119
6.5.4 Reconstruction of statistics . . . 121
6.6 The mean-field approximation . . . 125
6.7 Concluding remarks . . . 127

7 Conclusion 130
7.1 Summary of results . . . 130
7.2 The importance of Lagrangian information . . . 132
7.2.1 Galilean transformed mean and Rossby wave spectra . . . 132
7.2.2 Topology and coherent structures . . . 133

7.3 The importance of higher order Casimirs 134
7.4 Towards a statistical theory of the ocean mesoscale 136
7.4.1 Entropy production 136
7.4.2 Extension to more realistic flows 137
7.5 Concluding remarks 138

A Boundary condition for finite deformation radius **140**

B Galilean transformed mean and Rossby wave spectra **143**

C Numerical computation of entropy **148**

D Effect of stochastic white noise on eddy-mixing entropy **150**

E Time-evolution of maximal entropy **153**

F Significance testing **155**

G Reconstructing maximum entropy statistics **158**

Bibliography **160**

List of Figures

1.1 *Right* - set up of closed basin coupled with circumpolar channel. *Left* - plot of the response of mean circumpolar current to changed wind forcing. *Blue* - eddies are parameterized. *Red* - eddies are barely resolved. *Green* - eddies are partially resolved. Figure reproduced from Munday et al. (2013) 5
1.2 The blue box has enough eddies within to allow the mean to be statistically representative. The red box is close to the typical eddy size leaves the average within the box sensitive to the fluctuations of the turbulent flow. 6

2.1 Two surfaces of constant neutral density γ_n enclosing a circulation, Γ, about a closed loop. Figure modified from notes by M. McIntyre. . . 10
2.2 Geometry of the model domain, consisting of a zonal channel of meridional width L_y and zonal extent L_x. The wind-stress, τ, is zonal and varies only with the meridional coordinate, y. 19

3.1 Schematic of the Kolmogorov (1941) picture of turbulence. Energy is added at the stirring (forcing) length scale, small wavenumbers or large length scales, and is removed at the dissipation length scale. The dissipation length scale at high wavenumber, k_ν can be inferred from Kolmogorov theory using dimensional arguments. Between the stirring and dissipation scales we assume an inertial region where energy is transferred between wavenumbers at a rate ϵ. Image from Vallis (2006). 23
3.2 Schematic of the theory of two-dimensional turbulence. Similar to the Kolmogorov picture, but now with two dissipation length scales, one at small wavenumbers and the other at large wavenumbers, and two inertial ranges. We see a transfer of energy from the stirring length scale to large scale dissipation - the inverse cascade. Enstrophy is transferred from the stirring length scale to small scale dissipation - the forward cascade. Image from Vallis (2006). 26
3.3 Schema of two layer geostrophic turbulence. Wavenumber increasing left to right. Energy transfer is shown by solid arrows while enstrophy transfer is shown by dashed arrows. Image from Vallis (2006). k_0 and k_D represent the typical wavenumbers of the large and small scale dissipation respectively, while k_R is the typical wavenumber of barotropization. 28

3.4 Geometry of domain. wind stress, τ, in this study points from left to right. L_x and L_y are the longitudinal and latitudinal extents of the domain respectively. Macro-cells (blue) and micro-cells (red) in this study. Exploiting zonal symmetry, zonal bands are used as macro-cells whilst micro-cells are the grid-points within. 32

4.1 *Left*: examples of positively skewed (*red*) and negatively skewed (*blue*) distributions compared with a normal distribution (*black*). *Right*: examples of leptokurtic (*red*), kurtosis greater than three, and platykurtic (*blue*), kurtosis less than three, distributions compared with a normal distribution with a kurtosis of three (*black*). All distributions shown have a mean of 0 and a standard deviation of 1. 40

4.2 Snapshots of potential vorticity from the statistically steady state for simulations D and R. Simulation D, eastward wind stress, has strong cross-channel potential vorticity gradients forming a mixing barrier which meanders due to persistent travelling Rossby waves. Simulation R, westward wind stress, does not show the same behaviour but instead exhibits mixing across the channel. 44

4.3 Snapshots of potential vorticity from the statistically steady state for simulation U which has a westward wind forcing. Here we do not see a large-scale mixing barrier prohibiting mixing across the channel, but we do see mixing barriers wrapped up around two large coherent eddies. 45

4.4 Schematic of the various flow regimens simulated in this study as a function of the non-dimensional wind stress strength, τ_0, and linear drag coefficient, r. The regimes are classified by inspection where mixing barriers are defined as persistent, strong potential vorticity gradients; strong eddies are defined as persistent large eddies; and weak eddies are defined as smaller transient eddies. 46

4.5 a) and b) schema of the potential vorticity profile at the time of meeting the Rayleigh-Kuo criterion (4.2) for eastward and westward jets respectively. Dashed line: βy. c) and d) plots of the second derivative of the velocity profile at the time of meeting the Rayleigh-Kuo criterion; the dashed indicates the value β. Westward jets meet the Rayleigh criterion in the centre. Eastward jets meet the Rayleigh criterion in the flanks. In each case the black dots indicate where the criterion is met. 47

4.6 Time-averages for simulations A-G: a) time-averaged potential vorticity against meridional distance and; b) time-averaged streamfunction against meridional distance. Note, given a sufficiently long time the time-averages become zonally symmetric. 49

4.7 a) Mean potential vorticity against normalized time-averaged streamfunction. b) Standard deviation of potential vorticity against normalized time-averaged streamfunction. The streamfunction is normalized by it's steady state boundary value, $\overline{\Gamma}$, which is the integrated zonal momentum. Each plot shows the results from seven experiments each corresponding to a doubling of the forcing strength from the previous. The data from panels a and b are rescaled with respect to their norm to give; c) the average across all seven experiments of the rescaled mean potential vorticity as the solid-black curve; d) the average across all seven experiments of the rescaled standard deviation of the potential vorticity as the solid-black curve. In both c and d the dashed-red curves give the standard error. 50
4.8 The mean (a and c) and standard deviation (b and d) as functions of time-averaged streamfunction are shown for the dissipation (a and b) and reversed forcing (c and d) experiments. 52
4.9 Snapshot of steady state potential vorticity for simulation H, $r = 0.0008$. A large eddy with very negative potential vorticity is persistent in the bottom half of the domain. 53
4.10 a) skewness of potential vorticity as a function of time-averaged streamfunction. The line $S = 0$ is shown and represents statistics which is symmetric about its mean. b) kurtosis of potential vorticity as a function of time-averaged streamfunction. The lines $K = 3$ and $K = 1$ are shown. A kurtosis of 3 corresponds to a Gaussian distribution, a kurtosis greater than this is referred to as leptokurtic and while a kurtosis of greater than three is referred to as platykurtic. A kurtosis of 1 is the smallest possible kurtosis and corresponds to an unbiased coin toss. In c and d the skewness and the kurtosis are given respectively from the bimodal distribution, (4.10), and the trimodal distribution, (4.12). The parameters used are $d = 10$ and $c = 0.6$. 55
4.11 The skewness (a and c) and kurtosis (b and d) as functions of time-averaged streamfunction are shown for the dissipation (a and b) and reversed forcing (c and d) experiments. Again the $S = 0$ line is shown for the skewness and the $K = 3$ and $K = 1$ for the kurtosis. 59

4.12 This figure shows different ways to look at the statistics in the middle of the jet for forcing experiment D. *Top* - joint histogram of potential vorticity and streamfunction at the central meridional distance (close to the $\overline{\psi} = 0$ line). *Middle row* - three histograms of potential vorticity with the statistics sampled in different ways. *Bottom row* - streamfunction in various frames of reference are plotted illustrating the ways the statistics were sampled for the middle histograms. *Left column* - statistics taken from positions defined by the time-averaged streamfunction, $\overline{\psi} = 0$ contour, these positions are fixed in time. *Middle column* - statistics taken from positions defined by the Galilean transformed time-averaged streamfunction, $\overline{\psi_{GT}} = 0$ contour, these positions are also fixed in time. *Right column* - statistics taken from positions defined by the instantaneous streamfunction, $\psi = 0$ contour, these positions are allowed to change with time. The scale for the distribution above is omitted as it can be normalized away and we are only interested in the shape of the distribution at present. 62
4.13 a) the mean, b) standard deviation, c) skewness and d) kurtosis as functions of instantaneous streamfunction are show. Here the statistics has been sampled along contours of streamfunction for all times. . . . 63
4.14 Plots of the joint histograms of potential vorticity and streamfunction for the forcing experiment, simulations A-G. Colour scale gives the number of counts in each of the 100×100 bins. *Brighter - more counts. Black - zero counts. White - maximum counts.* 66

5.1 a) Time-mean surface relative vorticity and b) bathymetry (depth) of the Southern Ocean. The locations of geographical regions which are relevant to this study: A, the region of the Agulhas current retroflection (as it encounters the Antarctic Circumpolar Current) with Agulhas rigs propagating north-west from the region towards the South Atlantic; B, the region of the Kerguelen Plateau; C, the region of the Campbell Plateau; C, the approximate region of the DIMES experiment (Gille et al., 2007), passing through the Drake Passage; E, the region of the Malvinas western boundary current and its extension. 78
5.2 a) Skewness and b) kurtosis of surface relative vorticity (de-trended and with seasonal cycle removed) in the region of the Agulhas current retroflection for the whole time period. 80
5.3 a) A horizontal cross-section of potential vorticity kurtosis at depth $z = -328$m. b) A vertical cross-section of potential vorticity kurtosis at longitude of 143.4°. c) A vertical cross-section of potential vorticity kurtosis at longitude of 288.2°. d) A vertical cross-section of time-mean zonal velocity at longitude of 143.4°. e) A vertical cross-section of time-mean zonal velocity at longitude of 288.4°. f) A vertical cross-section of time-mean zonal potential vorticity at longitude of 143.4°. g) A vertical cross-section of time-mean zonal potential vorticity at longitude of 288.2°. 83

5.4 Detailed plot of high-low-high kurtosis feature in a vertical cross-section at longitude a) 143.4°and b) 288.2°. Contours of time-mean neutral density are shown. 84
5.5 Maps of skewness and kurtosis for two neutral density surfaces from Southern Ocean State Estimate data. a) Skewness at $\gamma_n = 26.78$. b) Kurtosis at $\gamma_n = 26.78$. c) Skewness at $\gamma_n = 27.21$. d) Kurtosis at $\gamma_n = 27.21$. Boxes indicate regions of interest. We can see positive-negative features in the skewness concomitant with high-low-high features in the kurtosis. These can be interpreted as meandering potential vorticity fronts corresponding to a mixing barrier. *White* - regions of no data due to outcropping or land. 86
5.6 a) A vertical cross-section in neutral density coordinates of skewness at longitude of 143.4°. b) A vertical cross-section in neutral density coordinates of kurtosis at longitude of 143.4°. c) A vertical cross-section in neutral density coordinates of skewness at longitude of 288.2°. d) A vertical cross-section in neutral density coordinates of kurtosis at longitude of 288.2°. 88
5.7 a) Typical snapshot of potential vorticity on neutral density surface, $\gamma_n = 27.2\mathrm{kg\,m^{-3}}$, during winter. b) Same is plotted during the subsequent summer (6 months later). c) Map of kurtosis where all points which outcrop at any time during the time-series have been masked out. Now, we do not see the longitudinally-extended high-low-high feature. 89
5.8 a) Map of the eddy enhancement factor ϵ on a neutral density surface. b) Map of kurtosis of potential vorticity on a neutral density surface. The neutral density surface for a and b is $\gamma_n = 27.2\mathrm{kg\,m^{-3}}$. 92
5.9 Global histograms of potential vorticity for the neutral density surfaces, $\gamma_n = (26.71, 26.96, 27.22, 27.48)\mathrm{kg\,m^{-3}}$. Histograms for the a) full statistics as well as b) the $z > -200$m and c) the $z < -200$m sub-populations. 93
5.10 Spatial distribution of potential vorticity and the corresponding probability distribution under strong diffusive mixing for region A, with prescribed boundary values, and region B, with no prescribed boundary values. 96
5.11 Joint histograms of potential vorticity and an ordinate for the neutral density surfaces, $\gamma_n = (26.71, 26.96, 27.22, 27.48)\mathrm{kg\,m^{-3}}$. a) Joint histogram, $n(Q, x|\gamma_n)$. b) Joint histogram, $n(Q, y|\gamma_n)$. c) Joint histogram, $n(Q, z|\gamma_n)$. 98

6.1 Entropy as a function of time for different values of linear drag coefficient, simulations $D_1, ..., D_7$. A - We find that exponential growth of barotropic instabilities lead to a very fast growth in the entropy. B - The entropy reaches a maximum value which is insensitive to the value of the linear drag coefficient. This is followed by a decrease in entropy at a rate greater than the contribution of the linear drag. This implies that this decrease is due to the eddies themselves. C - At long time the entropy decay is linear and is explained by the linear drag. For the lowest values of linear drag coefficient a persistent Rossby wave emerges causing a secondary increase in the entropy before the long-time behaviour is seen. 107
6.2 Measured slope of long time linear decay of the entropy, diagnosed from simulations $D_1, ..., D_7$, plotted against the predicted slope. The agreement is near perfect. 109
6.3 Evolution of entropy during spin-up of forced-dissipative flow. A - Initially laminar flow with small perturbation. As the flow becomes unstable the eddy-mixing entropy increases rapidly. B - Entropy reaches a global maximum at the point where turbulence has spread across the whole domain. Disorder is at small scales. C - Coincident with the emergence of a large scale Rossby wave the entropy decreases. Subsequently the entropy fluctuates around a balanced time-mean value. . 110
6.4 (a) Evolution of entropy in simulations $FD_1, ..., FD_7$. We see that entropy behaves as a balanced quantity in a statistically steady state: in the time mean the sources of entropy are equal to the sinks of entropy. (b) Evolution of energy in simulations $FD_1, ..., FD_7$ shown on a logarithmic scale. Like entropy the energy is balanced in a statistically steady state. The energy takes a longer time than the entropy to reach steady state balance. 112
6.5 Reconstruction of the entropy evolution using relation (6.12). *Top* - using only the first two Casimirs, circulation and enstrophy. *Bottom* - using first ten Casimirs. Both show significant correlation with the two Casimir reconstruction clearly matching the low frequency fluctuations, however but failing to capture the high frequency fluctuations. The ten Casimir reconstruction performs substantially better. 116
6.6 (a) Correlation between the diagnosed and reconstructed entropy time series following the relationship (6.12). Entropy is calculated on zonal bands for simulations $FD_1, ..., FD_7$. We see a clear improvement with number of Casimirs included. This improvement is quantitatively different for each simulation. (b) Correlation between the diagnosed and reconstructed entropy time series following the relationship (6.12). Entropy is calculated on instantaneous streamfunction for simulations $FD_1, ..., FD_7$. We see a clear improvement with number of Casimirs included. This improvement is quantitatively different for each simulation but now all simulations show high correlation of greater than 0.7 with only four Casimirs. This is to be contrasted with a). 117

6.7 (a) Shows the evolution of $\log_{10}\alpha$ with time. We see a very quick relaxation to a steady state value with fluctuation around this value in steady state being small. (b) Shows the steady state mean value of α against the strength of the wind stress. We see at low forcing have a very high sensitivity of entropy to energy while at high forcing the sensitivity is low. (c) Shows mean value of α against wind stress strength, now on a log-log scale. We see a near power-law behaviour with an approximate slope of -1.7. 122

6.8 Diagnostics testing reconstructed maximum entropy statistics. (a) Comparison of the time-averaged mean potential vorticity against time-averaged streamfunction for simulation FD_6. (b) Comparison of the time-averaged standard deviation of potential vorticity against time-averaged streamfunction for simulation FD_6. (c) Comparison of the time-averaged skewness of potential vorticity against time-averaged streamfunction for simulation FD_7. (d) Comparison of the time-averaged kurtosis of potential vorticity against time-averaged streamfunction for simulation FD_7. (e) Comparison of the maximum entropy probability distribution and the diagnosed distribution from simulation FD_4, distributions are evaluated for the centre of the channel, $y = 0$.(f) Comparison of the maximum entropy probability distribution and the diagnosed distribution from simulation FD_5, distributions are evaluated for the centre of the channel, $y = 0$. 124

B.1 Spectrum of root mean square difference between time mean and transformed time mean against test wavespeed for run A. We see peaks due to higher and lower wavenumber Rossby waves to either side of the dominant peak. 145

B.2 Spectrum of root mean square difference between time mean and transformed time mean against test wavespeed for run D. This is one of the messiest of the spectra with large peaks near to the most dominant. . 146

B.3 Spectrum of root mean square difference between time mean and transformed time mean against test wavespeed for run G. For lower wavenumbers the peaks have become broader and it becomes consequently easier to find the dominant Rossby wave. 147

F.1 *Red line* - mean cross-validated correlation between diagnosed and reconstructed entropy as a function of wind stress strength. *Black lines* - percentiles for the ensemble of mean cross-validated correlations from 'synthetic' time series of energy and Casimirs. 156

Chapter 1

Introduction

1.1 Motivation

Geostrophic turbulence is of interest to a wide range of scientific fields, including meteorology, oceanography, climate science and planetary science. Perhaps most strikingly, geostrophically turbulent flow is characterized by the plethora of large-scale and long-lived fluid-flow phenomena within the Earth's and atmosphere (e.g., stratospheric polar vortex, jet stream) and ocean (e.g., Kurishio and Gulf Stream extensions, Agulhas rings, Antarctic Circumpolar Current). Beyond our planet we find spectacular examples of the emergent phenomenology of geostrophic turbulence such as Saturn's polar hexagon (Godfrey, 1988), Jupiter's Great Red Spot (Hide, 1961) and now Jupiter's many other features which have been vividly exposed by the recent JUNO observations (Bolton et al., 2017; Kaspi et al., 2018). These phenomena have inspired equally varied and interesting approaches in the attempt to understand the emergent physics driving them, including: dynamical systems approaches to western boundary extensions (Dijkstra and Ghil, 2005); general relativity inspired interpretations of ocean rings (Haller and Beron-Vera, 2013); and statistical modeling of Saturn's hexagon (Farrell and Ioannou, 2017), to give only a few examples. These

fascinating, and often puzzling, phenomena continue to give cause for us to improve our fundamental theoretical description of geostrophic turbulence in order to understand these features in their own right as well as to understand their consequences for planetary weather and climate.

One of the distinguishing differences between oceanic and atmospheric geostrophic turbulence is the typical length scale of eddies. The length scales of eddies are set by the (baroclinic) deformation radius which is of order 40km for the ocean and of order 1000km for the atmosphere (e.g. Salmon, 1998). We see that accurately representing geostrophic turbulence within the atmospheric component of a comprehensive coupled climate models requires less computational power compared to the oceanic component. Currently, the eddies which make up ocean mesoscale (on scales of 10-100km) turbulence lie beyond the computational reach of modern climate models which typically have an ocean resolution of 100km (e.g. Fox-Kemper et al., 2014). Thus, it becomes necessary to parameterize the effect of eddies on the large scale flow.

An additional difficulty in understanding geostrophic turbulence arises from its complex interactions across a wide range of scales. As has been mentioned, geostrophic turbulence is characterized by its interaction with large scales but equally exhibits strong interactions with the smaller-scale ocean mixed layer whereby it can influence other components of the climate system such as ocean-atmosphere coupling and biological production (e.g. Lapeyre and Klein, 2006; Lapeyre et al., 2006; Lévy, 2008; Sasaki et al., 2014).

The problem of strong interactions between length and time scales is a generic feature within the highly complex climate system as a whole: there is no 'spectral gap'. That is, the climate system exhibits activity and interactions at all conceivable temporal and spatial scales, which requires innovative and sophisticated theoretical modeling approaches (e.g. Majda and Klein, 2003). Although geostrophic turbulence

is better resolved in the atmospheric component of climate models, compared to the ocean component, there remains computational limitations in representing other turbulent flows such as convection and clouds (e.g. Schneider et al., 2017). In addition, interactions occur between regimes governed by a wide range of physical, chemical and biological processes (Peixóto and Oort, 1984).

Although we will concentrate the ocean mesoscale within this thesis, it should be remembered that this is a somewhat arbitrary distinction in scales which is not cleanly respected by nature. Additionally, it should be noted that the theoretical nature of this work allows much of the intuition formed in this thesis to have relevance to geostrophically turbulent systems beyond the ocean.

1.2 Ocean mesoscale turbulence

Ocean mesoscale turbulence plays an important role in setting the dynamical and thermodynamical balances in the Earth's ocean basins. This makes geostrophic turbulence an important ingredient in understanding the global ocean circulation (e.g. Marshall et al., 2002); its feedback onto the climate system (e.g. Lovenduski et al., 2007; Munday et al., 2014); and understanding the effect of geostrophic turbulence on the global heat and carbon dioxide uptake of the ocean (e.g. Marshall and Zanna, 2014). Ocean jets such as the Gulf Stream, Kuroshio Extension, as well as the jets embedded within the Antarctic Circumpolar Current, represent some of the most turbulent regions of the ocean (e.g., Stammer and Wunsch, 1999). Such jets have an important influence on the large scale circulation of the ocean and atmosphere. For example, Scaife et al. (2011) highlight that model representation of winter atmospheric blocking is greatly improved when the horizontal resolution of the North Atlantic ocean is increased, thereby improving the representation of the Gulf Stream. From a hydrographic survey, Bower et al. (1985) determined that water masses are

mixed across the Gulf Stream at depth, but that the Gulf Stream acts as a barrier towards the surface where the jet is strongest, inhibiting exchange of tracers between the sub tropical and subpolar gyre.

The particular focus of this thesis will be the Southern Ocean where we also find turbulent ocean jets playing an important role in mediating ocean large scale mixing. Ferrari and Nikurashin (2010) and Klocker et al. (2012) show the suppression of mixing, and consequently the transport of heat and salt, across jets within the Antarctic Circumpolar Current. This suppression of mixing can be attributed to the tendency of strong jets to form mixing barriers, regions of high potential vorticity gradient which eddies have difficulty penetrating.

Moreover, mesoscale eddies plays a leading order role in setting the stratification and the transport of the Antarctic Circumpolar Current (Karsten et al., 2002). Mesoscale eddies are uniquely important in the Southern Ocean for several reasons. As discussed in Rintoul et al. (2001) turbulent eddies are required to explain the balance of heat fluxes in the Southern Ocean via a southwards eddy flux of heat. This balance is related to the momentum balance of the Antarctic Circumpolar Current, which is closed by eddy fluxes (Johnson and Bryden, 1989). In the Southern Ocean zonal momentum is input at the surface, due to wind stress, and removed at the bottom, due topographical form stress; consequently, eddies are needed to flux momentum from the surface to the abyss. One particular phenomenon associated with Southern Ocean eddies will be discussed in the following section and provides an example for why improved understanding of the ocean mesoscale is necessary.

1.2.1 The Southern Ocean and eddy saturation

A particular illustration of the importance of mesoscale eddies for the climate system is *eddy saturation* (Straub, 1993; Hallberg and Gnanadesikan, 2001; Tansley and Marshall, 2001; Hallberg and Gnanadesikan, 2006; Munday et al., 2013), where the

Antarctic Circumpolar Current volume transport does not change significantly with increased wind-stress strength. The study of Munday et al. (2013) looks at an idealized Atlantic/Southern Ocean system and shows clearly the cost of not explicitly resolving ocean eddies by considering the response of steady state transport to wind-forcing in simulations with varying levels of eddy representation. In Figure 1.1 the

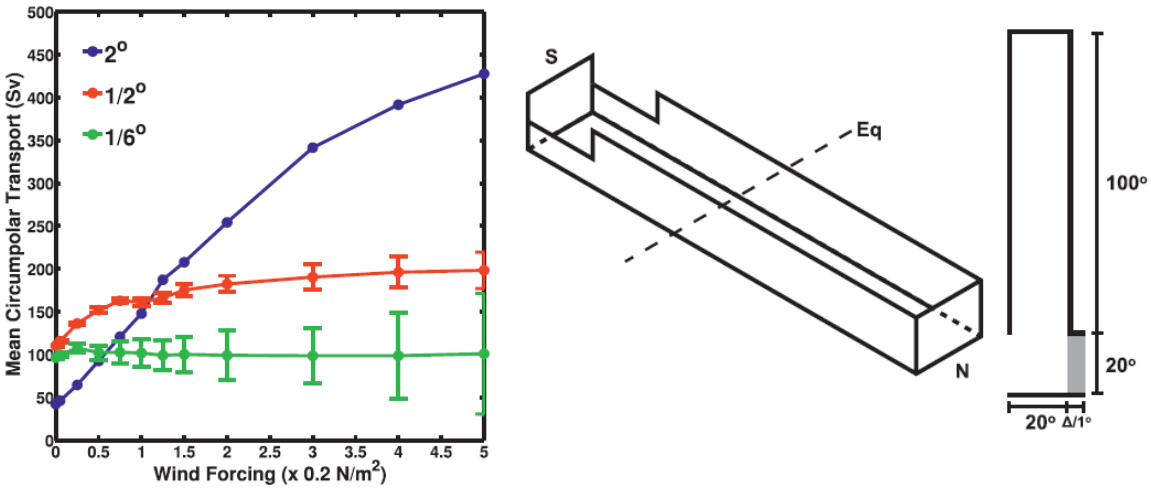


Figure 1.1: *Right* - set up of closed basin coupled with circumpolar channel. *Left* - plot of the response of mean circumpolar current to changed wind forcing. *Blue* - eddies are parameterized. *Red* - eddies are barely resolved. *Green* - eddies are partially resolved. Figure reproduced from Munday et al. (2013)

response of the Southern Ocean transport to increased wind forcing is given for three different resolutions: eddies not resolved (blue), eddies barely resolved (red) and eddies partly resolved (green). In the cases where the eddies are not resolved the eddies are parametrized by the usual Gent-McWilliams parameterization scheme (Gent and McWilliams, 1990; Gent et al., 1995). Figure 1.1 clearly shows that resolved eddies can change the mean flow's response to changed forcing as compared to parameterized eddies; the resolved eddies render the mean circumpolar transport insensitive to increased wind forcing whereas parameterized eddies exhibit a strong response. Behaviour such as this makes an improved understanding of geostrophic turbulence a key priority in considering the ocean's role in the complex climate system.

1.2.2 Stochastic Parameterization

The discussion in previous sections show that there is a need to improve the parameterized representations of mesoscale turbulence in ocean models. The ocean eddy parameterizations are examples of a turbulent closure. The turbulence closure problem, which comes from the non-linear advection term in the Navier-Stokes equation, is one of great unsolved problems in classical physics.

Currently, the parameterizations typically used in ocean and climate models are variants of that developed by Gent and McWilliams (1990) and Gent et al. (1995) at a coarse resolution. Recent studies have explored the possibility of a stochastic approach (Berloff, 2005; Porta Mana and Zanna, 2014; Grooms et al., 2015; Zanna et al., 2017) for ocean models. There are several motivations for adopting a stochastic approach.

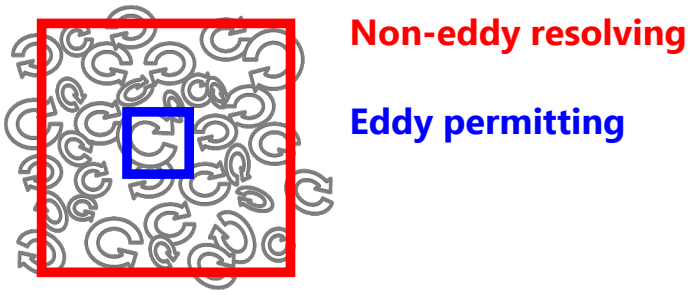


Figure 1.2: The blue box has enough eddies within to allow the mean to be statistically representative. The red box is close to the typical eddy size leaves the average within the box sensitive to the fluctuations of the turbulent flow.

- The fact that stochastic noise is added to the model provides a natural ensemble just by running the model several times (Buizza et al., 1999). This is useful in the context of climate prediction where measures of model uncertainty are difficult to obtain.

- It is known that the addition of noise to non-linear models can cause changes in their mean (Gardiner, 2004). In fluid models this can be used to mimic the

effect of small-scale and fast turbulent motions. Thus it is important to get the small scale variation right to ensure a correct large scale mean behaviour.

- As the resolution in models approaches the eddy permitting regime, we break the assumptions of time and length scale separation required for the valid use of a deterministic parameterization (e.g. Gent and McWilliams, 1990; Gent et al., 1995). For a deterministic parameterization to work, we need there to be many eddies contained in the grid-cell to ensure that a mean is representative of their behaviour. In the case where the grid-cell is near the typical size of an eddy (for example, the deformation radius) then the effect of the large scale flow will vary wildly as individual eddies pass through the grid-cell as schematically demonstrated in Figure 1.2.

In light of these advantages of describing and parameterizing eddies stochastically it is important to study ocean mesoscale turbulence from a statistical and theoretical perspective.

1.3 Approach of this thesis

The two examples discussed in this motivational chapter highlight the need for a more sophisticated statistical theory for the ocean mesoscale. In this thesis I will take an empirical approach, considering what ocean mesoscale turbulence statistics should look like and which statistical features are important to understand dynamically and to reproduce, as well as a theoretical approach where fundamental turbulence theories are considered.

In this thesis I will employ a combination of methodologies including idealized numerical simulations, comprehensive ocean models constrained by observational data as well as heuristic theoretical analysis. By employing these methodologies in conjunction with each other I am able to consider general organizing principles of geostropic

turbulence while also testing their validity in both idealized and comprehensive simulations.

The structure of this thesis is as follows. In Chapter 2, the relevant background from geophysical fluid mechanics is given and the numerical considerations for the barotropic channel model used in this study is presented. In Chapter 3, statistical theories of turbulence are briefly reviewed and their relevance to the ocean mesoscale is discussed. In Chapter 4, a study of the statistical nature of jets in an idealized barotropic channel ocean is presented and is also published in David et al. (2017). In Chapter 5, a study of potential vorticity statistics derived from the Southern Ocean State Estimate is presented exploring how the intuitions formed in Chapter 4 can be used to understand a more realistic ocean model. In Chapter 6, a study of an eddy-mixing entropy is presented and the utility of an equilibrium statistical mechanics theory of geophysical turbulence is considered in the context of a forced-dissipative flow. In Chapter 7, the results of this thesis are summarized, the common themes of Chapters 4, 5 and 6 are elucidated and areas worth consideration for future work are highlighted.

Chapter 2

Geophysical fluid dynamics background

2.1 Introduction

Geophysical fluid dynamics is the study of rotating and stratified fluid motions of planetary oceans and atmospheres. Geophysical fluid dynamical models, of varying degrees of approximation, range from the primitive equations, appropriate for modeling oceans and atmospheres in climate and weather models, to highly simplified quasi-geostropic models, appropriate as theoretical thinking tools for the study of geostrophic turbulence. In this chapter, the concepts and models of geophysical fluid dynamics relevant for this thesis are presented.

In Section 2.2, the concept of potential vorticity is introduced and its fundamental importance to geophysical fluid dynamics is highlighted. In Section 2.3, the quasi-geostrophic approximation is outlined and various quasi-geostrophic models are introduced. In Section 2.4, we examine the conserved quantities of ideal quasi-geostrophic dynamics. In Section 2.5, the Hamiltonian formulation of ideal quasi-geostrophic dynamics is explored. In Section 2.6, the particular geophysical fluid model used for

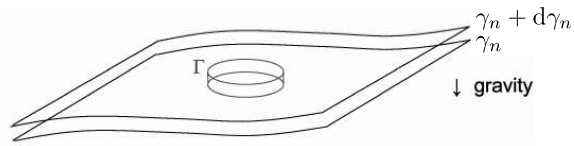


Figure 2.1: Two surfaces of constant neutral density γ_n enclosing a circulation, Γ, about a closed loop. Figure modified from notes by M. McIntyre.

much of this thesis is presented along with some numerical considerations. In Section 2.7, some results from geophysical fluid dynamics are presented with respect to the potential vorticity fronts, jets and mixing barriers.

2.2 Potential vorticity

On a fundamental level the material conservation of potential vorticity emerges from Noether's theorem of classical mechanics, corresponding to a fluid parcel relabeling symmetry of the Lagrangian (e.g. Salmon, 1998). This is an expression of the fact that the physics is invariant with respect to the arbitrary way we may choose to label fluid parcels. For example, we could 'name' all the fluid parcels according to their starting positions or we could equivalently use their finishing position. Potential vorticity conservation is an expression of Kelvin's circulation theorem combined with thermodynamics. Its generic validity means that potential vorticity conservation holds for a hierarchy of successively approximate ocean models.

In geophysical fluid mechanics, potential vorticity is a quantity proportional to the absolute circulation integrated between two isentropic[1] surfaces, see Figure 2.1, and is given by

$$Q = \frac{(\boldsymbol{\omega} + \mathbf{f}) \cdot \nabla \sigma}{\rho}, \tag{2.1}$$

where $\boldsymbol{\omega}$ is the three-dimensional relative vorticity; $\mathbf{f} = 2\boldsymbol{\Omega} \sin\theta$, is the Coriolis pa-

[1]This entropy here is the thermodynamic entropy of a fluid associated with molecular motions. It should not be confused with the eddy mixing entropy which I discuss in this thesis; this will be discussed in detail in Chapter 3

rameter; σ is an isentropic thermodynamic variable; and ρ is the in-situ density.

In other words, the potential vorticity is a projection of the absolute vorticity, that is the sum of the relative vorticity and the vorticity due to the Earth's rotation, onto the the 'thermodynamic vertical', the direction normal to a surface of constant thermodynamic entropy. Equivalently in an geophysical context, neutral density[2] (ocean) or potential temperature (atmosphere) can be used instead of entropy.

2.3 Quasi-geostrophic dynamics

In this thesis I will be mainly interested in fluid motions in which the dominant balance is geostrophic, thus making quasi-geostrophic dynamics an essential theoretical tool for studying turbulence in this regime. In quasi-geostrophic dynamics potential vorticity also plays an important role. However, the approximating assumptions must be understood in order to understand the difference between quasi-geostrophic potential vorticity and the aforementioned Ertel potential vorticity. The assumptions of quasi-geostrophic dynamics are given in Vallis (2006):

- The Rossby number is small. The Rossby number is a non-dimensional number which gives the relative importance of the advective term and the Coriolis term in the Navier-Stokes equation in a frame of reference rotating with angular velocity $\mathbf{\Omega}$. The Rossby number, Ro, is given by

$$Ro \sim \frac{|(\mathbf{u}\cdot\nabla)\mathbf{u}|}{|\mathbf{\Omega}\times\mathbf{u}|} \sim \frac{U}{f_0 L}, \tag{2.2}$$

  where, L and U are the characteristic length and velocity scales respectively, $f_0 = 2\Omega\sin\theta_0$ is the constant approximation to the Coriolis parameter and θ is the latitude.

[2]The details of neutral density will be further discussed in Chapter 5.

- The length scale is comparable to the deformation scale, $L_d = NH/f_0$ where, $N = \left(-\frac{g}{\gamma_n}\frac{\partial \gamma_n}{\partial z}\right)^{\frac{1}{2}}$ is the Brunt-Väisälä buoyancy frequency of the, stratification of the mean neutral density, γ_n; H is the characteristic depth; and f_0 is the constant approximation to the Coriolis parameter. Specifically we require that

$$\left(\frac{L}{L_d}\right)^2 Ro = \mathcal{O}(Ro). \tag{2.3}$$

- Variations in the stratification will be small compared to the background stratification.

- Variations in the Coriolis parameter are small. So if $f = f_0 + \beta y$ then $|\beta L|/|f_0| = O(Ro)$.[3]

- The characteristic time scale is given by $T = L/U$.

Once these approximations have been applied, the quasi-geostrophic potential vorticity, q, has a Lagrangian tendency given by

$$\frac{Dq}{Dt} = F + D, \tag{2.4}$$

where the Lagrangian (material) derivative is denoted as

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla, \tag{2.5}$$

and $\mathbf{u}$ is the geostrophic velocity of the flow, such that the streamfunction, ψ, is related to the velocity by $\mathbf{u} = \nabla \times (\psi \mathbf{k})$. The potential vorticity, q, takes different forms depending of the level of approximation applied and will be discussed below. On the right hand side we have dissipation, D, and forcing, F; when these are both

[3]The Coriolis parameter is given by $f = 2\Omega \sin(\theta)$, where Ω is the angular velocity of the earth and θ is the latitude. Assuming a spherical Earth we have, $\beta = 2\Omega \cos(\theta)/R_\oplus$, where $R_\oplus$ is the radius of the Earth.

zero we refer to the system as *ideal*. The simplicity of (2.4), an implicit equation in only one variable[4], q, and its ability to capture a lot of ocean and atmospheric dynamical behaviour makes the quasi-geostrophic equation an important theoretical tool in geophysical fluid dynamics.

The form that the potential vorticity takes depends on the level of approximation applied to the background stratification of the flow. Assuming a continuous background stratification, the potential vorticity takes the form

$$q = \nabla^2\psi + \beta y + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right) \tag{2.6}$$

where q and ψ are functions of vertical coordinate, z. The first term in the quasi-geostrophic potential vorticity is the relative vorticity of the flow; the second term is the variation in the Coriolis parameter, represented up to linear order by the term βy; and the third term is a vortex stretching term where $\frac{\partial \psi}{\partial z}$ can be interpreted as the buoyancy anomaly.

It should be noted here that quasi-geostrophic potential vorticity is materially conserved along depth surfaces, unlike Ertel potential vorticity which is materially conserved along isentropic surfaces. However, under the quasi-geostrophic approximation the background, or time-mean, isentropic surfaces are aligned with depth surfaces to leading order and anomalies in the stratification are assumed to be small; consequently, the physical interpretations of Ertel potential vorticity can, to a good approximation, be carried over to quasi-geostrophic potential vorticity in many situations. If necessary, it is also possible to formulate quasi-geostrophic dynamics in isentropic coordinates (Berrisford et al., 1993) and the two formulations, isopycnal surface and depth surface, yield approximately equal quasi-geostrophic potential vorticity. Additionally, it is of practical convenience that the gradient of Ertel potential vorticity on isopycnal surfaces are proportional to gradients of quasi-geostrophic po-

[4]That is, $\mathbf{u}$ and ψ can be derived diagnostically from knowledge of q.

tential vorticity along depth surfaces.

The continuously stratified case is in practice approximated by the multi-layer quasi-geostrophic potential vorticity

$$q_i = \nabla^2\psi_i + \beta y + \frac{f_0^2}{H_i}\left(\frac{\psi_{i-1}-\psi_i}{g'_{i-1}} - \frac{\psi_i - \psi_{i+1}}{g'_i}\right), \tag{2.7}$$

where; f_0 is the 0^{th} order approximation to the Coriolis parameter; $g'_i = g(\rho_{i+1} - \rho_i)/\rho_0$ is the reduced gravitational acceleration which takes into account the effect of buoyancy with ρ_0 being a reference density; and subscripts denote the layer which each variable is on.

When there is barotropic flow (e.g. for the shallow water approximation) the potential vorticity is further simplified to

$$q = \nabla^2\psi + \beta y - \frac{\psi}{L_{BT}^2}, \tag{2.8}$$

where $L_{BT} = \frac{\sqrt{gH}}{f}$ is the barotropic, or external, deformation radius. For the model used in this thesis we make another level of approximation and assume a rigid lid, removing the ageostrophic vortex stretching term, to obtain

$$q = \nabla^2\psi + \beta y. \tag{2.9}$$

The simplicity of this model makes it an attractive model for studying fundamental statistical theories of turbulence while retaining the form of potential vorticity material conservation, (2.4), shared by more complex models. The numerical considerations for solving this model along with the choices of the forcing, F, and dissipation, D, are given in Section 2.6.

An important distinction to made between barotropic and baroclinic quasi-geostrophic dynamics is that the former does not contain any available potential energy while the

latter does. This means that barotropic instabilities takes energy from the *kinetic* energy of the large-scale flow while baroclinic instabilities take energy from the *available potential* energy. Because of this distinction, baroclinic models (two-layers or more) can exhibit a Lorenz energy cycle which is a robust feature of realistic geophysical flow (Lorenz, 1955). Despite this phenomenological deficiency, studying barotropic flow remains an important starting point as the lack of baroclinic effects allows us to study eddy-mean flow interaction without the complications of having to understand potential and kinetic energy partitions. In addition, although ocean mesoscale turbulence is largely baroclinic in nature, barotropic instability plays an important role in the large scale dynamics of turbulent ocean jets (Waterman et al., 2011).

2.4 Invariants of motion

It is important to note that the ideal dynamics of barotropic flow conserves the following quantities. In ideal flow, within a closed domain, the energy

$$E = \frac{1}{2} \int \mathrm{d}^2\mathbf{x}\, (\nabla\psi) \cdot (\nabla\psi) \tag{2.10}$$

is conserved, where the integral is taken over the spatial domain of the fluid, representing a global conservation of an invariant of motion. Global conservation should not be confused with material conservation; for example, energy is described as being globally conserved while potential vorticity is described as being materially conserved.

In addition, ideal flow conserves the integral of any function of potential vorticity

$$C = \iint \mathrm{d}^2\mathbf{x}\, c(q), \tag{2.11}$$

where c is an arbitrary function. These quantities are called Casimirs (e.g. Salmon, 1998). We are primarily interested in the polynomial Casimirs which we will denote

as

$$C_n = \iint \mathrm{d}^2\mathbf{x}\, q^n. \tag{2.12}$$

Two Casimirs of particular physical importance are the circulation, $n = 1$, and the enstrophy, $n = 2$. Alternatively, all Casimirs can be conserved simultaneously by conserving the global potential vorticity distribution

$$\Pi(q) = \frac{1}{\mathcal{A}} \frac{dA(q)}{dq} \tag{2.13}$$

where $A(q)$ is the area of the domain occupied by points with a value of potential vorticity less than q, and is therefore a cumulative distribution function; and $\mathcal{A}$ is the total area of the domain.

2.5 Hamiltonian formalism

In the absence of forcing and dissipation, quasi-geostrophic dynamics can be re-cast under a Hamiltonian formalism, where the equations of motion are written in terms of the energy (or Hamiltonian) of the system. We include a discussion here as it is an important foundation of the statistical mechanics theory introduced in Chapter 3 as well as giving an alternative interpretation of the conservation of the Casimir invariants of motion.

Hamiltonian mechanics is usually applied to single or multi-particle systems. There are subtle complications when this formalism is extended to continuous media. This is the subject of the final chapter of Salmon (1998) and a review of Hamiltonian fluid mechanics is given in Morrison (1998). The Hamiltonian form of barotropic quasi-geostrophic dynamics (Zakharov and Piterbarg, 1988, and references therein) is given by

$$\frac{\partial q}{\partial t} = \{q, H\}_{NC}, \tag{2.14}$$

where H is given by

$$H[q] = -\frac{1}{2}\int \psi(q - \beta y)\mathrm{d}\mathbf{x} \tag{2.15}$$

which is the energy, as given in Section 2.4. In fact the evolution of an arbitrary functional of q, for example, $A[q]$ is given by

$$\frac{\partial A}{\partial t} = \{A, H\}_{NC}. \tag{2.16}$$

The bracket above is known as the non-canonical Poisson bracket and is defined by

$$\{F, G\}_{NC} = \int q \left\{ \frac{\delta F}{\delta q}, \frac{\delta G}{\delta q} \right\} \mathrm{d}\mathbf{x}, \tag{2.17}$$

where, $\{A, B\} = \partial_x A \partial_y B - \partial_x B \partial_y A$. The term 'non-canonical' here refers to the fact that the formalism has changed from that used for particle dynamics where the Poisson bracket satisfies three conditions:

- antisymmetry, $\{A, B\} = -\{A, B\}$,
- Jacobi identity, $\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0$,
- non-singularity, $\{A, B\} \neq 0$,

where A, B and C are independent, non-constant functionals. In the non-canonical case the non-singularity condition has been lifted and now we can find functionals, $C[q]$, such that

$$\{C, A\}_{NC} = 0, \tag{2.18}$$

for all $A[q]$. One such A would be the Hamiltonian itself and thus by (2.16) these Cs are conserved quantities. These turn out to be the Casimirs given in (2.12) and (2.13). Understanding the Hamiltonian structure of these objects underlies the work described in Section 3.3 and the implications of this Hamiltonian structure on the

phase space of two-dimensional flow is discussed in Zeitlin (1991). The Hamiltonian formalism for the baroclinic case is presented in Holm (1986).

2.6 The barotropic channel model

The non-dimensional, prognostic equation to be solved, in Chapters 4 and 6, is

$$\frac{\partial q}{\partial t} = -\{\psi, q\} - r\nabla^2\psi - \nu_h\nabla^6\psi - \partial_y\tau(y), \tag{2.19}$$

where ψ is the streamfunction; $\tau(y)$ is the zonal wind stress which varies only with meridional coordinate y, with a maximum strength given τ_0; r is the linear drag coefficient; and ν_h is the hyper-viscosity. Although we will refer to τ as the wind stress, τ is, more correctly, the wind stress normalized by density and depth. Linear drag is included to remove momentum at large scales and the hyper-viscosity is included in order to preferentially damp enstrophy at small scales. The potential vorticity is given by

$$q = \nabla^2\psi + \beta y, \tag{2.20}$$

where $\nabla^2\psi$ is the relative vorticity and βy is the planetary vorticity. In this study we refer to q as the potential vorticity, and to the global integral of q^2 as the enstrophy, in order to facilitate comparison with studies that incorporate baroclinicity; however, in the barotropic model the potential vorticity is identical to the absolute vorticity.

The non-dimensional parameters of the above model are given in terms of their

dimensional counterparts, denoted by $*$, as

$$\nu_h = \frac{\nu_h^*}{UL^3}, \tag{2.21}$$

$$\beta = \frac{\beta^* L^2}{U}, \tag{2.22}$$

$$r = \frac{r^* L}{U}, \tag{2.23}$$

$$\tau_0 = \frac{\tau_0^* L}{U^2}, \tag{2.24}$$

where, U and L are a characteristic velocity and length scale respectively.

The numerical code used is a modified version of that developed for the freely decaying turbulence problem in Esler (2008a,b) and, originally, Esler and Haynes (1999). The model is solved using a quasi-spectral method, decomposing into Fourier modes in the zonal direction making use of the periodicity. In the meridional direction a finite difference method is used. The first obvious difference is that the original code solved for two-layer quasi-geostrophic dynamics whereas the modified code solves for a barotropic flow with a rigid lid. Secondly, a wind stress forcing has been added which required an alternative boundary condition to be applied to ensure the solution is consistent with momentum conservation. The domain geometry is given in Figure 2.2. The boundary conditions follow those specified in McWilliams (1977) requiring

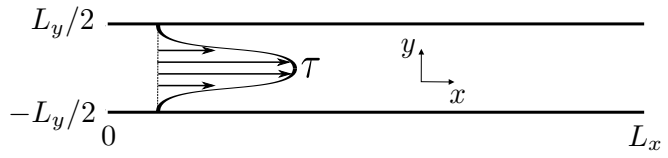


Figure 2.2: Geometry of the model domain, consisting of a zonal channel of meridional width L_y and zonal extent L_x. The wind-stress, τ, is zonal and varies only with the meridional coordinate, y.

an integral momentum balance to be applied at each time-step. The no normal flow condition is written

$$\partial_x \psi|_{N,S} = 0, \tag{2.25}$$

where the subscripts refer to the northern and southern boundaries respectively. To calculate the hyper-viscosity term we need two boundary conditions. To ensure no boundary stress we use the free slip conditions,

$$\nabla^{2n}\psi|_{N,S} = 0, \tag{2.26}$$

where $n = 1, 2$, ensuring that the stress at the boundary is zero and that the hyper-viscous term also does not produce a boundary moment sink. Finally, the momentum conserving condition is written

$$\psi|_{N,S} = \pm\frac{\Gamma(t)}{2}, \tag{2.27}$$

where $\Gamma(t)$ is spatially constant along the north and south boundaries and evolves temporally through the integral momentum balance,

$$\frac{d\Gamma}{dt} = -\int \mathrm{d}^2\mathbf{x}\left[r\frac{\partial\psi}{\partial y} + \tau\right], \tag{2.28}$$

which is exactly the form of boundary condition used in Treguier (1989). For the equivalent boundary condition for a barotropic channel with a finite deformation radius (no rigid lid approximation), see Appendix A.

2.7 Dynamics of jets, potential vorticity steps and mixing barriers

Having considered idealized models of geophysical fluid dynamics we now consider how some of the intuition developed for these models inform our knowledge of ocean turbulence. One of the important features of ocean mesoscale turbulence, mentioned in Section 1.2, is its contribution to the suppression of mixing. This suppression

of mixing can be attributed to the tendency of strong jets to form mixing barriers, regions of high potential vorticity gradient which eddies have difficulty penetrating.

The tendency for strong jets to form mixing barriers is, in fact, a consequence of the turbulence which sharpens the jets and maintain strong potential vorticity gradients by providing up-gradient fluxes of momentum (Starr, 1968). Evidence for this jet sharpening behaviour has a long history in baroclinic (e.g. Fultz et al., 1959) and barotropic (e.g. Sommeria et al., 1989) jets which is comprehensively reviewed in Dritschel and McIntyre (2008). The highly anisotropic picture we have formed from these studies is that of a potential vorticity step corresponding to the jet itself, sharpened by the eddies, with regions of homogenized potential vorticity to either side of the jet within which the eddies act more diffusively.

However, as we shall see in Chapters 4 and 5, the potential vorticity 'staircase' paradigm discussed above is too restrictive to describe the wealth of regimes of behaviour encountered in idealized models and the ocean alike. In Chapter 4, the distinction will be made between jets which exhibit mixing barrier behaviour and jets that do not. In Chapter 5, a further distinction is needed to separate out potential vorticity fronts which are not concomitant with mixing barriers. For these reasons we define the usage in this thesis of the following terminology.

- *Potential vorticity front:* is a region of high potential vorticity gradient separating regions with differing potential vorticity.

- *Mixing barrier:* is a potential vorticity front which inhibits mixing across the front.

- *Jet:* is a region of high mean velocity, often well aligned with the zonal direction.

Although it is clear that these three features of oceanic flow are linked it is important to distinguish them to ensure clear discussion.

Chapter 3

Statistical theories of geostrophic turbulence

3.1 Introduction

In this chapter I briefly introduce the theoretical concepts relevant to the study of ocean mesoscale turbulence. In Section 3.2, the homogeneous turbulence paradigm is introduced and its limitations for studying large scale ocean turbulence are discussed. In Section 3.3, an equilibrium statistical mechanics (Miller-Robert-Sommeria theory) of two-dimensional and geophysical turbulence is introduced. Although Miller-Robert-Sommeria theory does not suffer from all of the inhibiting assumptions of homogeneous turbulence theory, it is not clear how applicable Miller-Robert-Sommeria theory is in the context of forced-dissipative flow. In Section 3.4, the difference between a thermodynamic and information theoretic interpretation of entropy is discussed. In Section 3.5, the various applications of statistical physics and information theory to climate science are discussed.

3.2 The homogeneous turbulence paradigm

3.2.1 Three-dimensional turbulence

To start our discussion of turbulence the three-dimensional theory of Kolmogorov (1941) is considered, which is arguably the most famous result in the theory of turbulence. The prediction of Kolmogorov (1941) is that under a certain set of assumptions, the energy spectrum of three-dimensional turbulence can be written as

$$E(k) = C\epsilon^{\frac{2}{3}}k^{-\frac{5}{3}}, \tag{3.1}$$

where ϵ is the rate of energy dissipation per unit volume; $k = |\mathbf{k}|$ is the magnitude of the wavenumber; and C is Kolmogorov's universal constant. Figure 3.1 gives a visual

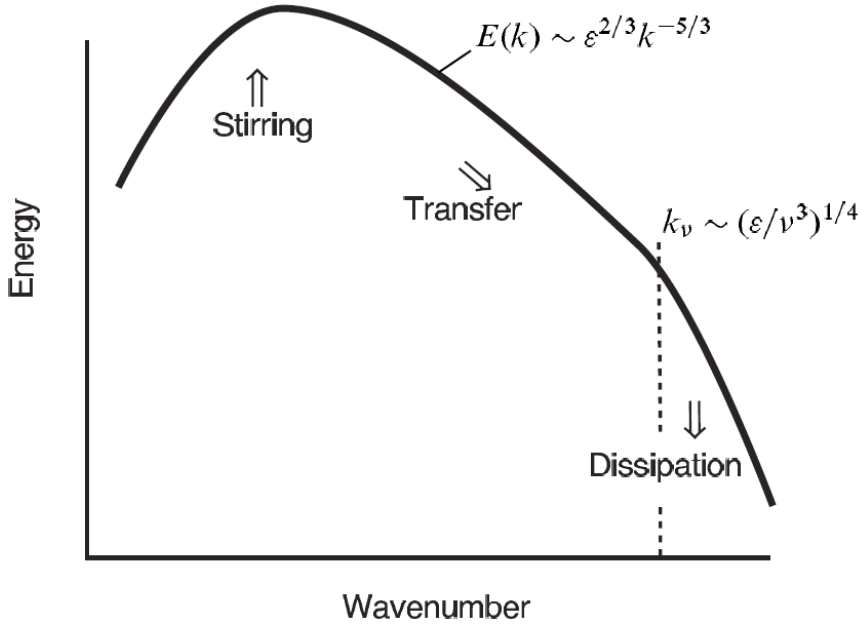


Figure 3.1: Schematic of the Kolmogorov (1941) picture of turbulence. Energy is added at the stirring (forcing) length scale, small wavenumbers or large length scales, and is removed at the dissipation length scale. The dissipation length scale at high wavenumber, k_ν can be inferred from Kolmogorov theory using dimensional arguments. Between the stirring and dissipation scales we assume an inertial region where energy is transferred between wavenumbers at a rate ϵ. Image from Vallis (2006).

description of the theory. To understand this equation let us examine its assumptions.

- *Incompressible fluid flow*, $\nabla \cdot \mathbf{u}$.

- *Homogeneous and isotropic turbulence.* Homogeneity allows us to think about the turbulence through Fourier methods which require an infinite domain, while writing $E \equiv E(k)$ only makes sense if the turbulence is isotropic, and hence radially symmetric in Fourier-space. Without isotropy E becomes a function of $\mathbf{k} = (k_x, k_y, k_z)$.

- *Inertial range.* It is assumed that there exists a range of length scales (or wavenumber) for which the fluid does not feel the effect of forcing (stirring) or dissipation. The picture here is that forcing acts at large length scales, dissipation acts at small length scales and has no affect at intermediate length scales, which are known as the *inertial range*.

- *Turbulent cascade.* It is also assumed that energy is transferred between length scales within the inertial range via non-linear interactions. Mathematically these are triad interactions between Fourier modes which become apparent when the non-linear term of the equation of motion is represented in Fourier-space. Physically these triad interactions can be thought of as eddies colliding and producing subsequent eddies. Inherent to the cascade is the idea that only eddies of similar sizes can interact and thus the transfer of energy becomes a *local process in Fourier-space.* Although it is common to use the terminology 'cascade' to mean a transfer of energy (or other quantity) between any scales, I will use the term 'cascade' to formally mean a local transfer in Fourier space throughout this thesis.

Utilizing these assumptions, Kolmogorov was able to proceed using simple dimensional arguments to derive (3.1). Kolmogorov reasoned that the energy spectrum should only depend on the rate of energy transfer due to the cascade, ϵ, which is a constant within the inertial range. Therefore, given the dimensions of the relevant

variables

$$[E(k)] = L^3T^{-2}, \tag{3.2}$$

$$[\epsilon] = L^2T^{-3}, \tag{3.3}$$

$$[k] = L^{-1}, \tag{3.4}$$

$$[\nu] = L^2T^{-1}, \tag{3.5}$$

the only possible form of the energy spectrum is given by (3.1), and the characteristic length scale where dissipation becomes dominant is given by the Kolmogorov scale

$$k_\nu = \epsilon^{\frac{1}{4}}\nu^{-\frac{3}{4}}. \tag{3.6}$$

A major inconsistency in Kolmogorov (1941) was pointed out in a famous footnote of Landau and Lifshitz (1987). Simply stated (e.g. Salmon, 1998): if we have two equal-volume turbulent systems which exhibit a Kolmogorov spectrum, each with a different dissipation rate, that is

$$E_1(k) = C\epsilon_1^{\frac{2}{3}}k^{-\frac{5}{3}}, \tag{3.7}$$

$$E_2(k) = C\epsilon_2^{\frac{2}{3}}k^{-\frac{5}{3}}, \tag{3.8}$$

then a composite (averaged) system with dissipation rate

$$\epsilon = \frac{1}{2}(\epsilon_1 + \epsilon_2), \tag{3.9}$$

and energy spectrum

$$E(k) = \frac{1}{2}(E_1(k) + E_2(k)), \tag{3.10}$$

does not itself have a Kolmogorov spectrum, that is

$$E(k) \neq C\epsilon^{\frac{2}{3}}k^{-\frac{5}{3}}. \tag{3.11}$$

Although there have been experimental results consistent with the Kolmogorov spectrum, as well as theoretical attempts to circumnavigate the Landau objection, Kolmogorov (1941) cannot be considered an universal theory of turbulence.

3.2.2 Two-dimensional turbulence

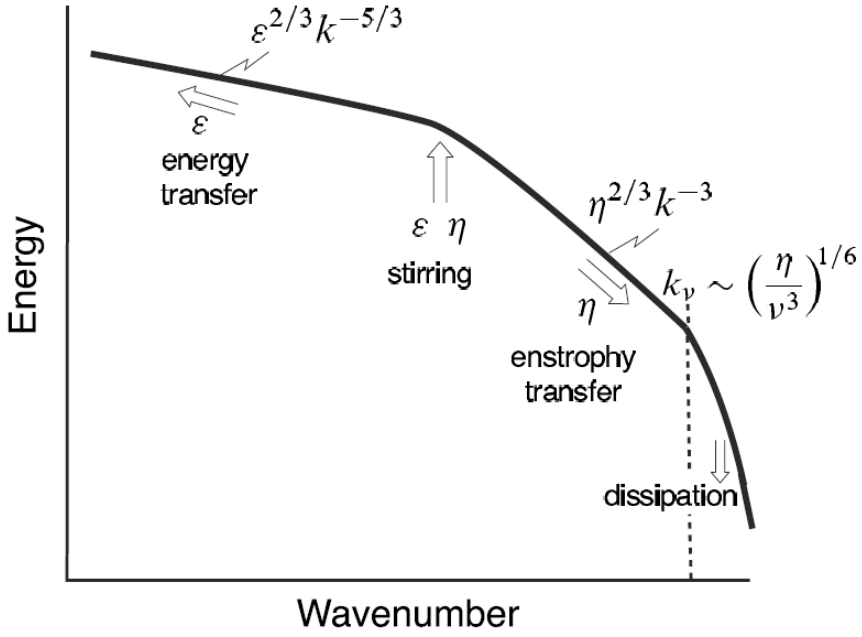


Figure 3.2: Schematic of the theory of two-dimensional turbulence. Similar to the Kolmogorov picture, but now with two dissipation length scales, one at small wavenumbers and the other at large wavenumbers, and two inertial ranges. We see a transfer of energy from the stirring length scale to large scale dissipation - the inverse cascade. Enstrophy is transferred from the stirring length scale to small scale dissipation - the forward cascade. Image from Vallis (2006).

Similar arguments can be made in the case of two-dimensional turbulence (Kraichnan, 1967) for which the assumptions stated in the previous sections are adopted.

Since in two-dimensions we have the additional conserved quantity [1] of enstrophy the picture becomes more complicated. Enstrophy is given by

$$C_2 = \int \mathrm{d}k\, k^2 E(k), \tag{3.12}$$

represented in Fourier-space. If we force the system at a wavenumber k_f then because of the additional constraint of having to conserve enstrophy we can deduce that no energy can be transferred to small scales in the limit of small viscosity. This implies that the energy is transferred to wavenumbers smaller than k_f (larger length scales). Thus, to have a steady state turbulence we must have a second source of dissipation at large length scales. Following this reasoning, we have the situation of two turbulent cascades a *direct* cascade of *enstrophy* to small scales with an energy spectrum,

$$E(k) = C_1 \eta^{\frac{2}{3}} k^{-3}, \tag{3.13}$$

and an *inverse* cascade of *energy* to large scales with an energy spectrum,

$$E(k) = C_2 \epsilon^{\frac{2}{3}} k^{-\frac{5}{3}}. \tag{3.14}$$

Here, ϵ is the energy transfer rate per unit volume as before and η is the enstrophy transfer rate per unit volume, see figure 3.2. The above results rely on the same assumptions as the three-dimensional case in addition to the added assumption of the secondary source of dissipation at large length scales. Again, as in the case of three-dimensional turbulence, the theory for two-dimensional turbulence contains an internal inconsistency. It can be shown (Salmon, 1998) that an energy spectrum of the form (3.13) necessarily breaks the assumption of locality in Fourier-space. As with the three-dimensional case, by modifying the theory this issue can be surpassed

[1] As was discussed in Chapter 2, there are in fact an infinite number of additional conserved quantities: a subject which we will return to in this thesis, Chapter 7.

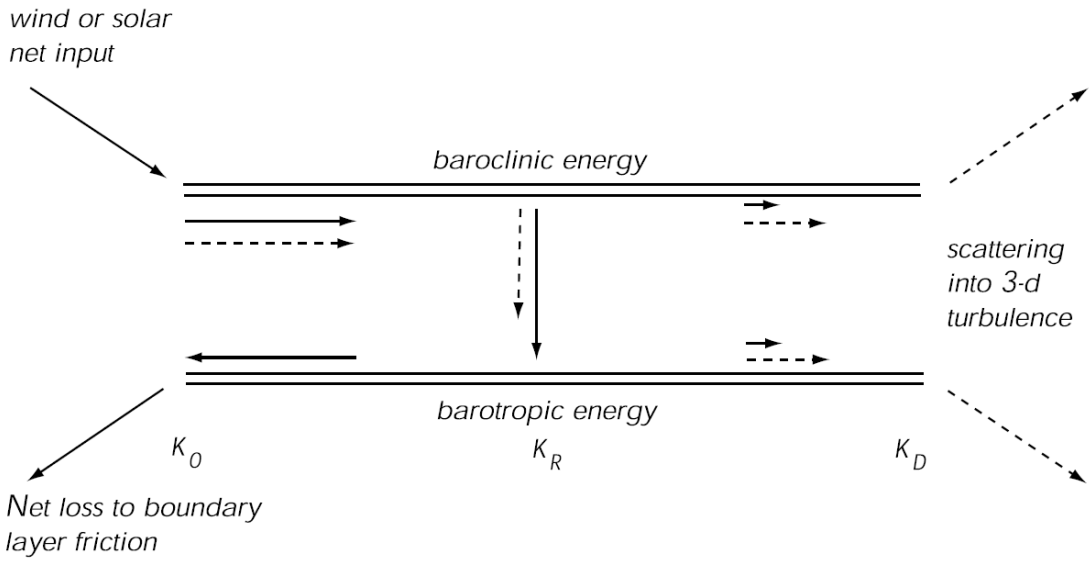


Figure 3.3: Schema of two layer geostrophic turbulence. Wavenumber increasing left to right. Energy transfer is shown by solid arrows while enstrophy transfer is shown by dashed arrows. Image from Vallis (2006). k_0 and k_D represent the typical wavenumbers of the large and small scale dissipation respectively, while k_R is the typical wavenumber of barotropization.

but its details go beyond the scope of this discussion.

3.2.3 Application to the geophysical flows

Figure 3.3 gives a picture of how the work of Charney (1971) is applied to geostrophic turbulence in a quasi-geostrophic model with multiple baroclinic modes. By looking at how a quasi-geostrophic model can be decoupled into barotropic and baroclinic modes, Charney (1971) claimed that baroclinic modes behave like the three-dimensional homogeneous turbulence while the barotropic mode behaves like the two-dimensional homogeneous turbulence. The baroclinic mode is forced by the wind at large length scales and energy cascades down to small length scales to be dissipated. The barotropic mode is forced by a process of *barotropization* where energy is transferred to the barotropic mode at length scales around the deformation radius. Energy is then transferred to large scales and is thought to be then dissipated by boundary layer friction, while enstrophy is transferred to small length scales to be dissipated by

viscosity. For a more detailed discussion the reader is referred to Vallis (2006) and Salmon (1998).

It is not clear how applicable the homogeneous turbulence paradigm is to the study of geophysical flows. Apart from incompressibility, which is a good approximation for geophysical flows, all the other assumptions presented in Section 3.2.1 are manifestly violated in the ocean mesoscale: the ocean mesoscale is highly inhomogeneous and anisotropic; with atmospheric forcing acting at a very wide range of time and length scales it hard to argue for the existence of an inertial range; and there is evidence that geostropic turbulence exhibits a direct transfer of energy from small to largest length scales (e.g. Nazarenko, 2011). In addition the study of Tung and Welch (2001) has found serious flaws in the arguments of Charney (1971), including a reliance on the result of Fjørtoft (1953) which is now known to be erroneous (Merilees and Warn, 1975). This calls into question the paradigm represented in Figure 3.3.

For the rest of this thesis I will follow a way of thinking about fluid turbulence which moves away from this turbulent cascade picture. This will involve an equilibrium statistical mechanics approach which allows us to lift the assumptions of homogeneity and isotropy as well as the assumption of a turbulent cascade.

3.3 Miller-Robert-Sommeria statistical mechanics

3.3.1 Statistical mechanics and turbulence

The first example of the application of statistical mechanics to two-dimensional turbulence comes from Onsager (1949), who proposed a model of singular point vortices. The statistical mechanics of point vortices has received much attention since then. Studies by Kraichnan (1967) and Kraichnan and Montgomery (1980), led the way towards continuous vorticity fields through utilizing the invariance of energy and enstrophy in variational problems. Geostrophic flows over topography were tackled

using this methodology: in Salmon et al. (1976) via the maximization of an entropy; and independently in Bretherton and Haidvogel (1976) via a phenomenological minimum enstrophy principle. Other contributions are given in Carnevale et al. (1981) and Carnevale (1982) which gives an alternative framework for the discussion of entropy evolution by calculating the entropy in terms of the energy spectrum of discretized dynamics.

In the early 1990s, work was published which led to a theory of equilibrium statistical mechanics of two-dimensional and geophysical flows (Miller, 1990; Robert, 1990, 1991; Robert and Sommeria, 1991; Miller et al., 1992). We refer to this theory as the 'Miller-Robert-Sommeria statistical mechanics'. Miller-Robert-Sommeria theory produces an equilibrium statistical mechanics (variational problem for finding the most probable macro-state of a closed system) of two-dimensional or quasi-two-dimensional flow which relies on the underlying Hamiltonian structure of the dynamics; applies to continuous vorticity fields; and conserves the mean-field energy and *all* Casimirs of the flow.

Reviews of Miller-Robert-Sommeria statistical mechanics are given in Sommeria (2001), Majda and Wang (2006), Chavanis (2009) and Bouchet and Venaille (2012). The strength of the Miller-Robert-Sommeria theory is that the work of Onsager (1949), Kraichnan and Montgomery (1980), Salmon et al. (1976) and Bretherton and Haidvogel (1976) are contained within this framework as particular limits or simplifications. This equilibrium/ideal (no forcing, no dissipation) theory has been used to suggest a statistical mechanical explanation for: the formation of ocean rings and jets (Venaille and Bouchet, 2011); the dynamics of the stratospheric polar vortex (Prieto and Schubert, 2001; Yasuda et al., 2017); Jupiter's Great Red Spot (Turkington et al., 2001; Bouchet and Sommeria, 2002); bottom trapped ocean currents (Venaille, 2012); the vertical structure in stratified quasi-geostrophic flow (Merryfield, 1998; Schecter, 2003; Venaille et al., 2012); as well as the global ocean circulation and

its associated density profiles is presented by Salmon (2012).

3.3.2 Eddy-mixing entropy

The entropy used in Miller-Robert-Sommeria statistical mechanics, the *eddy-mixing entropy*, is not the same as the thermodynamic entropy associated with molecular motions. The eddy-mixing entropy is a measure of the disorder of the large scale turbulent flow, and depends on the choice of coarse-graining which distinguishes between the large scales and the small scales of the flow. To formally define the eddy-mixing entropy we follow the approach presented in Jung et al. (2006). To proceed we will define two sub-systems of the full flow.

- A *micro-cell* which is the smallest scale over which the details of the flow are important. The micro-cell is equivalent to the grid-cell for a high resolution numerical simulation. We think of each micro-cell as being characterized by only one value of the potential vorticity.

- A *macro-cell* which is comprised of a number of micro-cells and is related to a choice of some coarse graining scale. The macro-cells should be chosen to exploit some dynamical symmetry of the system. In our case we, for the most part, choose zonal bands exploiting the zonal symmetries of the barotropic channel model used apart from in Section 6.5.2 where we choose contours of instantaneous streamfunction.

The macro- and micro-cells used in this study are schematically illustrated in Figure 3.4.

Having defined the macro-cells as such, it is possible to define an eddy-mixing entropy by counting the number of ways to arrange the micro-cells of a particular

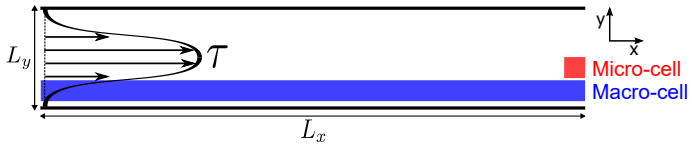


Figure 3.4: Geometry of domain. wind stress, τ, in this study points from left to right. L_x and L_y are the longitudinal and latitudinal extents of the domain respectively. Macro-cells (blue) and micro-cells (red) in this study. Exploiting zonal symmetry, zonal bands are used as macro-cells whilst micro-cells are the grid-points within.

binned value of potential vorticity into the macro-cells. This gives the expression

$$S = \ln W = \ln \prod_I \frac{M^{(I)}!}{\prod_r M_r^{(I)}!} \tag{3.15}$$

for the eddy-mixing entropy, where $M^{(I)}$ is the number of micro-cells in the I^{th} macro-cell and $M_r^{(I)}$ is the number of micro-cells with the r^{th} value of potential-vorticity bin in the I^{th} macro-cell. This counting method is adapted from Lynden-Bell (1967) where it was used to study the statistical mechanics of stellar systems.

For large numbers of micro- and macro-cells we can take the continuous limit to get

$$S[\rho] = -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\rho(\tilde{q}|\mathbf{x})\ln(\rho(\tilde{q}|\mathbf{x})), \tag{3.16}$$

in terms of the probability distribution function, ρ. In words, ρ *is the probability of measuring a value,* $\tilde{q}$*, of the potential vorticity at the point* $\mathbf{x}$ *in the domain.* It is important to note that in this expression $\mathbf{x}$ has taken the place of I in labeling the macro-cell. The coordinate $\mathbf{x}$ should be interpreted as a coarse-grained or smoothed coordinate. The $\tilde{q}$ is not the potential vorticity field, $q(\mathbf{x})$, but a random variable representing the result of a measurement of potential vorticity. It is important to note the difference between the probability distribution, $\rho(q|\mathbf{x})$, used here and the full probability distribution of the system, $p[q(\mathbf{x})]$.[2] We interpret ρ as a marginal distribution of p and they only become equal when neighbouring macro-cells become

[2]Square brackets denote a functional here.

statistically independent, which is the case for the *ideal* theory; this is discussed in more detail in Section 6.6. The eddy-mixing entropy is the sum over the continuous information entropies associated with the distribution of potential vorticity in each macro-cell. The method of numerically determining these entropies is given in Appendix C and is used throughout this study.

3.3.3 Maximization of entropy

The main result of Miller-Robert-Sommeria theory is the mean-field variational problem given by

$$\frac{\delta S}{\delta \rho} + \alpha \frac{\delta}{\delta \rho}\left(-\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho - E\right) - \frac{\delta}{\delta \rho}\left(\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\gamma(\tilde{q})\rho - \Pi(\tilde{q})\right) = 0, \quad (3.17)$$

for the maximization of entropy given the constraints imposed by the invariants of motion E and Π. Here α and $\gamma(q)$ are Lagrange multipliers and can be interpreted as the sensitivities of the entropy, S, to the energy, E, and global potential vorticity distribution, Π, respectively. The averaging operator $\langle\rangle$ is defined as an average over a macro-cell

$$\langle f(\tilde{q})\rangle = \int \mathrm{d}\tilde{q}\, f(\tilde{q})\rho(\tilde{q}|\mathbf{x}), \quad (3.18)$$

and leads to spatially-dependent averages. The variational problem, (3.17), is analogous to the variational problem used to derive the Boltzmann distribution in the statistical mechanics of gases. However, for molecular dynamics energy is the only relevant invariant. The variational problem uses the mean-field energy

$$E = -\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\,\langle\psi\rangle\langle\tilde{q}\rangle, \quad (3.19)$$

which is often used as an approximation of condensed matter physics. In Miller-Robert-Sommeria equilibria, however, the mean-field energy becomes equal to the

exact energy due to the non-local nature of the vortex interactions. The form of the variational problem presented here has been rigorously justified using large-deviation theory by Boucher et al. (1999).

The distribution which solves (3.17) is given by

$$\rho(\tilde{q}|\mathbf{x}) = \frac{1}{\mathcal{Z}} \exp\left(-\alpha\langle\psi\rangle\tilde{q} - \gamma(\tilde{q})\right), \tag{3.20}$$

where $\mathcal{Z}$ is the partition function. The spatial dependence of the distribution comes in through the average of the streamfunction defined through the solution of the Poisson equation

$$\nabla^2\langle\psi\rangle = \langle\tilde{q}\rangle - \beta y. \tag{3.21}$$

Equilibria which are the solution of (3.17) are characterized by a functional dependence between the averaged potential vorticity and the averaged streamfunction

$$\langle\tilde{q}\rangle = f(\alpha\langle\psi\rangle). \tag{3.22}$$

The function f can be proved to be monotonic and is related to the partition function through

$$f(\alpha\langle\psi\rangle) = \frac{\delta}{\delta(\alpha\langle\psi\rangle)} \ln\mathcal{Z}. \tag{3.23}$$

In the special case where we are only interested in systems which only conserve energy and enstrophy, often argued in the case of discretized dynamics, the equilibrium distribution becomes

$$\rho(\tilde{q}|\mathbf{x}) = \frac{1}{\mathcal{Z}} \exp\left(-\alpha\langle\psi\rangle\tilde{q} - \tilde{q}^2\right), \tag{3.24}$$

where γ_2 is now the Lagrange multiplier corresponding to the enstrophy C_2; the function, f, becomes linear

$$\langle\tilde{q}\rangle = -\frac{\alpha}{2\gamma_2}\langle\psi\rangle, \tag{3.25}$$

and the variance of potential vorticity becomes

$$\langle q^2 \rangle - \langle q \rangle^2 = \frac{1}{2\gamma_2}. \tag{3.26}$$

We present this special case here as the study of Jung et al. (2006) reproduce this result in a rotating annulus experiment of barotropic flow. One of the aims of Chapter 4 is to attempt to verify this result in a numerical simulation, and to test its sensitivity to forcing. In addition, numerical tests of Miller-Robert-Sommeria theory have been attempted (e.g. Wang and Vallis, 1994; Qi and Marston, 2014; Dritschel et al., 2015), however, it is difficult to produce a numerical model which is representative of *ideal* dynamics. In Chapter 6, the question of the utility of the equilibrium statistical mechanics approach in forced-dissipative models will be considered.

3.4 Thermodynamic versus information entropy

In Bayesian statistics and information theory the principle of maximum entropy states that given prior information about a system, the probability distribution function which best represents our state of knowledge is the one for which the *Shannon information entropy*

$$S = -\sum_i \rho_i \ln \rho_i \tag{3.27}$$

is maximal, where ρ_i is a discrete probability distribution function. Information entropy was introduced in Shannon (1948) and the maximum entropy method was first introduced in Jaynes (1957a) and Jaynes (1957b). The maximum entropy method is an important method of Bayesian statistics and information theory. The information entropy shares a close link with the *Gibbs entropy* which shares the same expression, (3.27), up to a constant of proportionality. Both the Shannon entropy and the Gibbs entropy can be interpreted as measures of uncertainty or information content.

In physics there exists an alternative interpretation of entropy and its relationship with the aforementioned interpretation it not always clear. In this interpretation the entropy is considered to be a measure of disorder and is given by the *Boltzmann entropy*

$$S = k_B \ln W, \tag{3.28}$$

where W is the number of possible microscopic configurations of the system which would yield the same macroscopic observables. The leap from the Boltzmann entropy to the Gibbs entropy can be easily made by interpreting the probability distribution, ρ_i, as a frequential probability. However, the Shannon entropy should be interpreted (and often, so is the Gibbs entropy) according to Bayesian statistics. This leads to confusion over understanding concepts such as the second law of thermodynamics. A recent study which highlights these issues and proposes a formalism which resolves the tension between the two interpretations, uncertainty versus disorder, is Bartolotta et al. (2015) where a Bayesian formulation of the Second Law of Thermodynamics is proposed. In this thesis I will use the two interpretations of entropy interchangeably and thus in various sections I will talk of the eddy-mixing entropy as being both a measure of disorder and as a measure of uncertainty or information.

3.5 Other applications of entropy and statistical mechanics to climate problems

At this stage it is appropriate to consider how the statistical mechanics presented above differs from various other applications of statistical physics ideas in climate science.

One of the ways entropic thinking has been used to understand the climate system is by consideration of the Maximum Entropy Production Principle (MEPP) which is

reviewed in Martyushev and Seleznev (2006). An attractive hypothesis for dealing with systems far from equilibrium, the MEPP often leads to controversy over the many fields of science where it has been applied. In climate science, its use to understand latitudinal temperature variational has had some success but fails to take into account important dynamical factors as it is a purely thermodynamical theory (e.g. Paltridge, 1975).

Whereas the MEPP theory of climate uses a thermodynamical entropy, information entropy has proved an useful tool for a variety of problems across an immense range of scientific fields, from the study of ant language (Ryabko and Reznikova, 2009), to quantifying the production of information about the climate from ice cores (Garland et al., 2016), to providing a measure of predictability in the context of climate modeling (Schneider and Griffies, 1999).

Statistical physics also gives a formalism for predicting the response of systems in equilibrium to small external forcing, the fluctuation-dissipation theorem (Nyquist, 1928), and its application for the climate has been proposed by Leith (1975). Since then the fluctuation-dissipation theorem has been considered for studying geophysical fluid dynamics (e.g. Lutsko et al., 2015) and various climate problems (e.g. Gritsun and Branstator, 2007; Ring and Plumb, 2008). The fluctuation-dissipation theorem, suffers from the assumption that systems are perturbed from an equilibrium state, which is not applicable to the climate system in general. Recently a generalized form of linear response theory has been derived (Ruelle, 1998, 2009), much like the fluctuation-dissipation theorem but applicable far from equilibrium. This theory can be used to understand the climate response to forcing and provides an invaluable tool for investigating possible climatic tipping-points (e.g. a transition to a snowball earth) (Lucarini and Sarno, 2010; Lucarini et al., 2010).

Having seen the various applications of statistical physics to climate science it is possible to classify them in the following ways.

- Applications of equilibrium and non-equilibrium thermodynamics at a planetary scale (e.g. Paltridge, 1975).

- Applications of information theory, for example, time-series analysis and predictability (e.g. Schneider and Griffies, 1999).

- Statistical mechanics of stochastic or chaotic dynamical systems, where the climate is interpreted as a high-order dynamical system (e.g. Lucarini et al., 2010).

- Statistical mechanics of turbulence, where a statistical mechanics approach is used to describe the turbulent motions of a continuous fluid (Miller-Robert-Sommeria statistical mechanics).

The final approach is the one followed throughout this thesis.

Chapter 4

The statistical nature of barotropic ocean jets

4.1 Introduction

This chapter presents the published results of David et al. (2017). The goal of this study is to understand the statistics of forced-dissipative ocean jets and how these are influenced by the dynamics of a mixing barrier. We explore the sensitivity of the jet statistics to forcing and dissipation and show that self-similarity exists across a wide range of parameters, describing the characteristic statistical signature of a mixing barrier.

To address this goal, we examine the mean state of the flow and its variability in the presence of forcing and dissipation using mean, standard deviation and higher order moments such as skewness and kurtosis. We will also compare and contrast these statistics to the predictions of equilibrium statistical mechanics. The energy and enstrophy conserving theory considered in this study is that presented in Jung et al. (2006). Jung et al. (2006) differs from other tests of equilibrium statistical mechanics (e.g. Qi and Marston, 2014; Dritschel et al., 2015) in that it derives a spatially

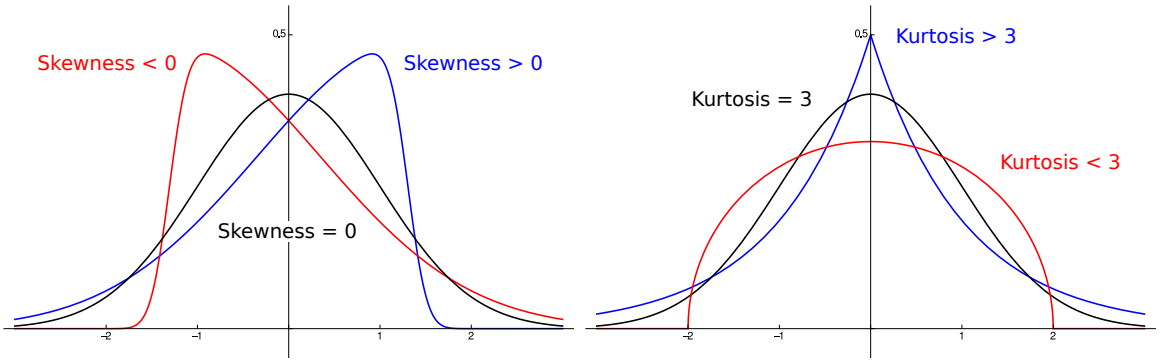


Figure 4.1: *Left*: examples of positively skewed (*red*) and negatively skewed (*blue*) distributions compared with a normal distribution (*black*). *Right*: examples of leptokurtic (*red*), kurtosis greater than three, and platykurtic (*blue*), kurtosis less than three, distributions compared with a normal distribution with a kurtosis of three (*black*). All distributions shown have a mean of 0 and a standard deviation of 1.

coarse-grained version of the theory in order to test the theory in a barotropic rotating annulus laboratory experiment. When only energy and enstrophy are conserved, the relationship between mean potential vorticity and the mean streamfunction is linear, (3.25), while the standard deviation is constant, (3.26), with respect to streamfunction according to equilibrium statistical mechanics. These predictions are derived in Jung et al. (2006) and are consistent with other interpretations of the theory (e.g. Salmon et al., 1976). Understanding these relationships in a forced-dissipative system, however, remains an open problem which we will discuss during the course of this study.

In addition to understanding how the mean and standard deviation relations change in the presence of forcing and dissipation, it is important to consider higher order moments of the flow, such as skewness and kurtosis, which exist only when the potential vorticity distribution is non-Gaussian. As illustrated schematically in Figure 4.1, different values of skewness and kurtosis of the probability distribution characterize asymmetry and intermittency (extreme events) in turbulent flow. These moments have been shown to be important in describing the statistics of passive tracers (e.g., Bourlioux and Majda, 2002). The character and nature of skewness

and kurtosis in turbulent geophysical flows have been considered by Thompson and Demirov (2006) and Hughes et al. (2010). In particular, Hughes et al. (2010) establish a mechanistic link between skewness and kurtosis and the large scale dynamics of jets, allowing these authors to use these moments to identify mixing barriers. In a highly non-linear dynamical system such as the ocean, we might expect extreme events in active tracers such as potential vorticity to have an impact on the mean large-scale circulation and to be sensitive to forcing and dissipation.

In this chapter, we study the potential vorticity statistics of a barotropic jet in a reentrant channel, varying the strength of forcing, the linear drag coefficient, and the direction of forcing. These experiments are used to pursue the following aims:

- To understand the dependence of the statistics of a turbulent flow on forcing and dissipation in the presence and absence of a mixing barrier.
- To establish a statistical picture of mixing barriers with a simple statistical model which will inform the stochastic parameterization of eddy-mean flow interaction.
- To distinguish between the statistics of eddies and the statistics of waves.

The results are restricted here to barotropic flows, an understanding of which is a prerequisite for understanding more complex flows with baroclinicity.

The structure of the chapter is as follows. In Section 4.2, the numerical experiments are presented. In Section 4.3, the dynamics of the west/eastward jet asymmetry is discussed. In Section 4.4, the mean and standard deviation of potential vorticity are calculated as a function of time-averaged streamfunction in order to contrast with the equilibria of the ideal dynamics, and a self-similar response to forcing and dissipation is discussed. In Section 4.5, the higher order moments of skewness and kurtosis are similarly considered. A clear link is made between the presence of a meandering mixing barrier and a characteristic skewness/kurtosis behaviour. In Section 4.6,

the importance of the meandering mixing barrier in determining the statistics of the flow is further considered, and the statistics calculated on instantaneous contours of streamfunction in order to follow the large scale meanderings of the mixing barrier in a quasi-Lagrangian sense. In Section 4.7, there is a further discussion of: joint potential vorticity and streamfunction statistics; the importance of mixed distributions; and implications for stochastic parameterization. Finally, Section 4.8 contains some concluding remarks.

4.2 Numerical experiments

For the domain geometry given in Figure 2.2, two parameters are varied across a wide range of experiments, labeled A through to U. In the first group of experiments, A-G, the strength of forcing, τ_0, is changed holding the linear drag coefficient, r, constant. The meridional variation of the wind stress is kept the same for all experiments,

$$\tau = \tau_0 \text{sech}^2 \left(\frac{y}{\delta} \right) \tag{4.1}$$

where δ is a constant controlling the width of the wind forcing. In the second group of experiments, H-N, the linear drag coefficient, r, is changed while holding the forcing strength, τ_0, constant. Finally the third group of experiments is a repeat of the first but with the direction of the forcing reversed. The parameters for all three experiments are given in Table 4.1 with the parameters which are held constant for all experiments are given in Table 4.2.

4.3 Linear stability and west/east asymmetry

During this study we will refer to jets as any peaked zonal mean flow whereas the term mixing barrier will refer to any region of high potential vorticity gradient which con-

simulation ID	Wind stress strength, τ_0	Linear drag coefficient, r	Timestep, dt
A	0.005	0.005	0.001
B	0.01	0.005	0.001
C	0.02	0.005	0.001
D	0.04	0.005	0.001
E	0.08	0.005	0.001
F	0.16	0.005	0.001
G	0.32	0.005	0.001
H	0.08	0.0008	0.0005
I	0.08	0.0009	0.0005
J	0.08	0.001	0.001
K	0.08	0.002	0.001
L	0.08	0.003	0.001
M	0.08	0.004	0.001
N	0.08	0.005	0.001
O	-0.005	0.005	0.001
P	-0.01	0.005	0.001
Q	-0.02	0.005	0.001
R	-0.04	0.005	0.001
S	-0.08	0.005	0.001
T	-0.16	0.005	0.001
U	-0.32	0.005	0.001

Table 4.1: List of experiments and non-dimensional parameters. Note that simulations N and E are identical.

Parameter	Value
Meridional extent, L_y	$5\pi/2$
Zonal extent, L_x	20π
Number of zonal grid-points, M	1024
Number of meridional grid-points, N	128
Time-step, dt	1×10^{-3}
Output frequency	1
Total time of output	1×10^{4}
Hyper-viscosity, ν_h	2×10^{-6}
Wind stress width parameter, δ	0.4
Beta parameter, β	0.2

Table 4.2: Fixed parameters of model simulations.

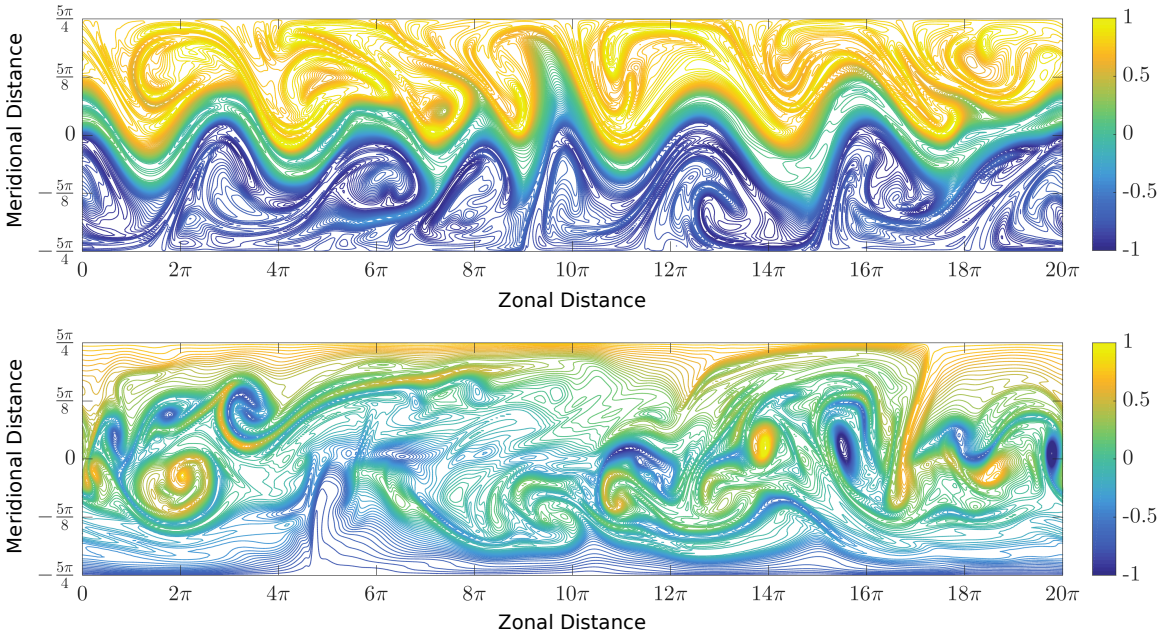


Figure 4.2: Snapshots of potential vorticity from the statistically steady state for simulations D and R. Simulation D, eastward wind stress, has strong cross-channel potential vorticity gradients forming a mixing barrier which meanders due to persistent travelling Rossby waves. Simulation R, westward wind stress, does not show the same behaviour but instead exhibits mixing across the channel.

sequently prohibits transport across it. A major dynamical difference is seen between the jets in simulations A-G (with varying forcing strength) and the jets in simulations O-U, with reversed forcing. This difference is evident in the snapshots of potential vorticity for simulations D and R, shown in Figure 4.2 for illustrative purposes as the flow is typical for the respective experiments: A-G and O-U. The main difference to observe is the clear loss of the narrow region of high cross-flow potential-vorticity gradient along with the concomitant large scale Rossby wave. This is a persistent structure in the eastward jet case for all forcing strengths but is not at all apparent in the westward propagating jet. It would not be correct to say that mixing barriers do not form at all in the reversed forcing experiments, this can be seen in a snapshot from simulation U in Figure 4.3. While there is no barrier prohibiting mixing across the channel, we do see mixing barriers which wrap themselves around two large coherent eddies, partially isolating potential vorticity within the two eddies. The topologically

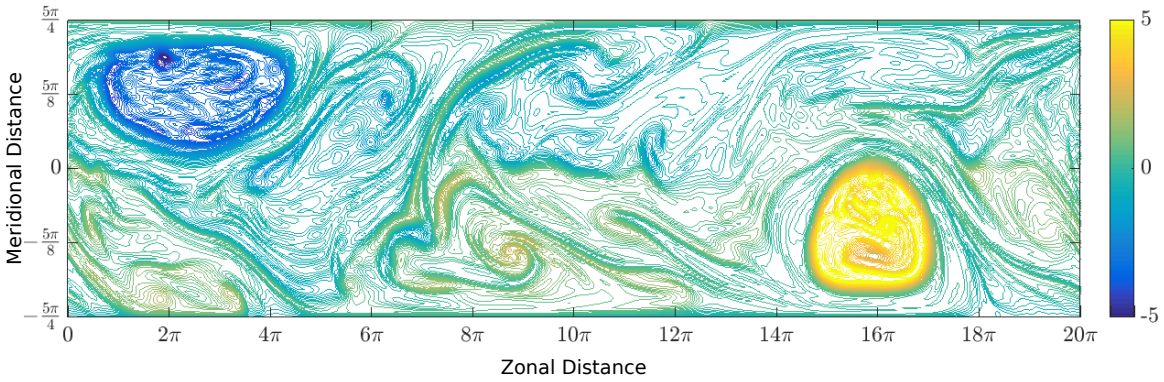


Figure 4.3: Snapshots of potential vorticity from the statistically steady state for simulation U which has a westward wind forcing. Here we do not see a large-scale mixing barrier prohibiting mixing across the channel, but we do see mixing barriers wrapped up around two large coherent eddies.

distinct nature of these mixing barriers changes the barriers' influence on the statistics as a function of time-averaged streamfunction, which is zonally symmetric. A schematic representing the various flow regimes simulated in this study is given in Figure 4.4.

This drastic difference between the two experiments implies that the dynamics are not symmetric with respect to a change in the sign of forcing. It can be seen from equations (2.19)-(2.20) that changing the sign of forcing is equivalent to changing the sign of β combined with a reflection in the meridional (y) axis. Making such a transformation leads to a drastic difference in the resulting statistically steady state flow. The presence and absence of this mixing barrier will form a large part of the discussion in the following sections.

The Rayleigh-Kuo inflection-point criterion (Rayleigh, 1880; Kuo, 1949) states that a necessary condition for barotropic instability of a zonal flow is that the meridional potential vorticity gradient changes sign, i.e.,

$$\beta - \frac{\partial^2 U(y)}{\partial y^2} = 0 \tag{4.2}$$

somewhere in the domain where $U(y)$ is the zonally symmetric velocity profile of

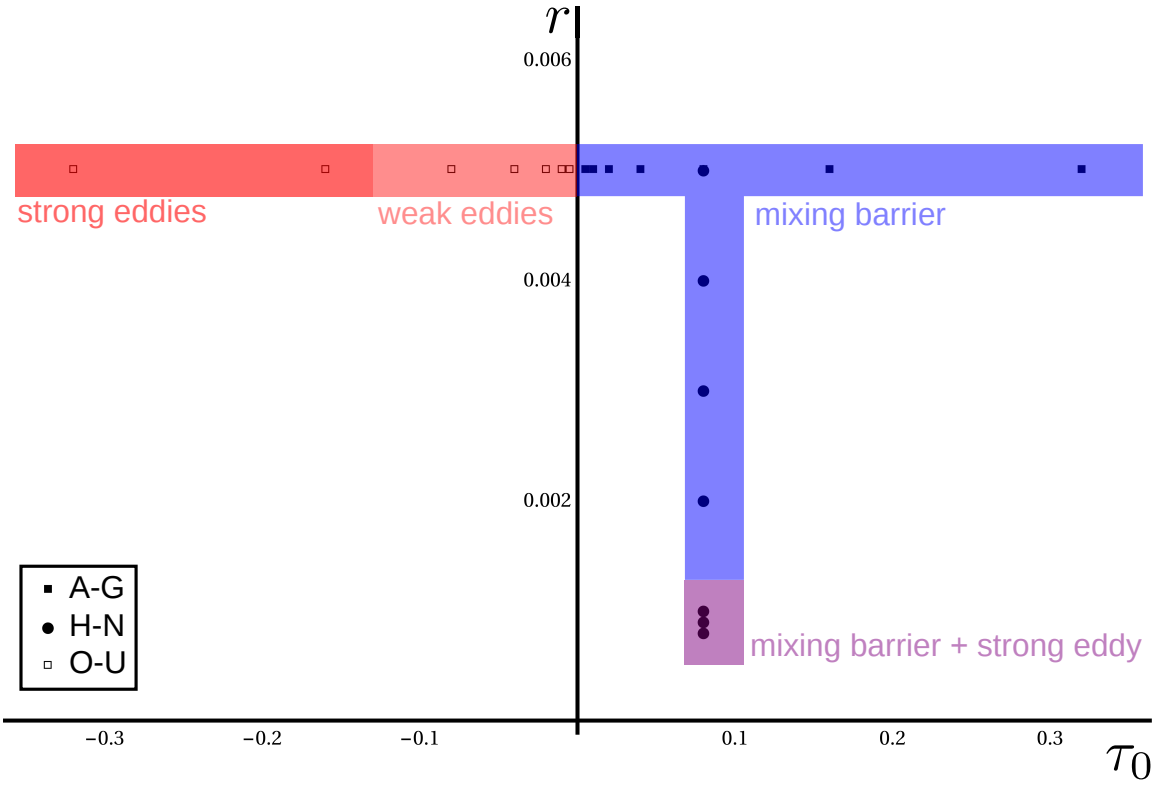


Figure 4.4: Schematic of the various flow regimens simulated in this study as a function of the non-dimensional wind stress strength, τ_0, and linear drag coefficient, r. The regimes are classified by inspection where mixing barriers are defined as persistent, strong potential vorticity gradients; strong eddies are defined as persistent large eddies; and weak eddies are defined as smaller transient eddies.

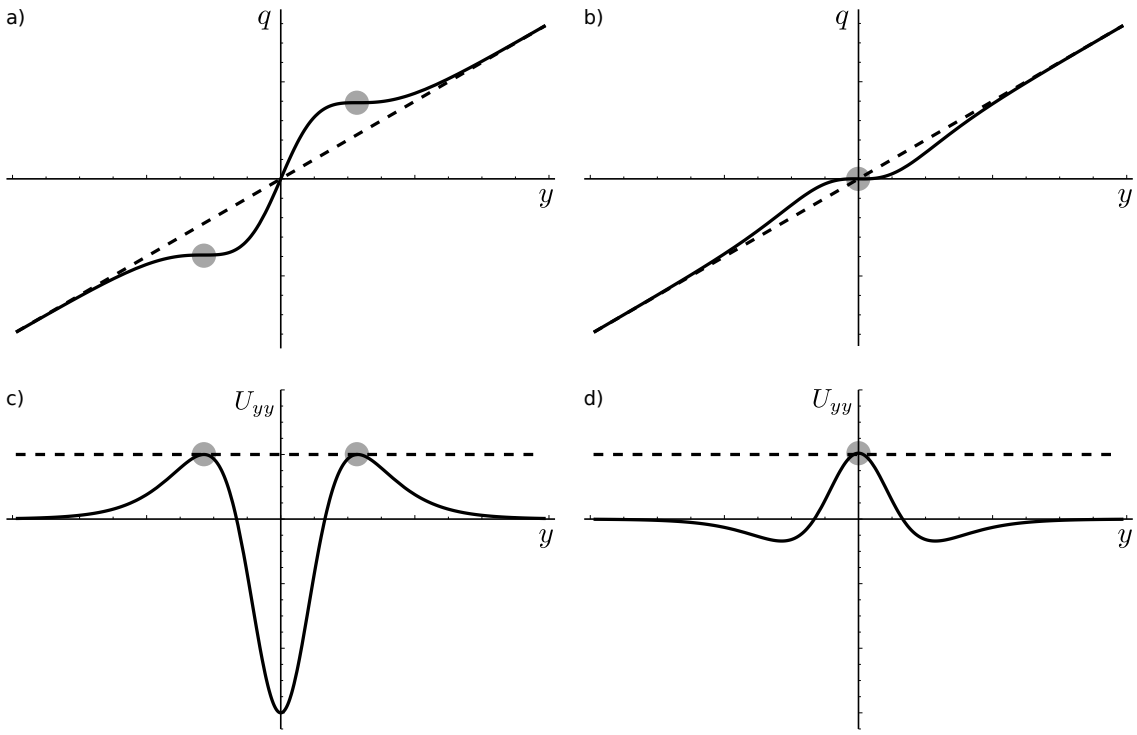


Figure 4.5: a) and b) schema of the potential vorticity profile at the time of meeting the Rayleigh-Kuo criterion (4.2) for eastward and westward jets respectively. Dashed line: βy. c) and d) plots of the second derivative of the velocity profile at the time of meeting the Rayleigh-Kuo criterion; the dashed indicates the value β. Westward jets meet the Rayleigh criterion in the centre. Eastward jets meet the Rayleigh criterion in the flanks. In each case the black dots indicate where the criterion is met.

the mean jet. While (4.2) is a necessary global condition for instability, in practice, instability often occurs in the locality that (4.2) is met (e.g., Waterman and Jayne, 2012; Waterman and Lilly, 2015). During spin up, the zonally symmetric forcing profile adds to the zonally symmetric flow until (4.2) is satisfied. Until this condition is satisfied the flow remains linear and the total potential vorticity input over time, t, is given by $-\partial_y \tau(y)t$, and corresponding momentum input of $\tau(y)t$. Figure 4.5 shows the potential vorticity profile and the second derivative of U at the time where Rayleigh-Kuo criterion is just met. There are two points to note here. Firstly, we can see analytically that the westward jet reaches the Rayleigh-Kuo criterion first, that is, (4.2) is satisfied for smaller τt in a westward jet. Secondly, the westward jet first

satisfies the Rayleigh-Kuo criterion (4.2) in the centre of the channel ($y = 0$) whereas the eastward jet satisfies the Rayleigh-Kuo criterion on its flanks, with the background potential vorticity gradient being reinforced at the centre of the eastward jet, through an up-gradient potential vorticity flux associated with the homogenization by the instabilities either side (Dritschel and McIntyre, 2008). It is these very same regions in which the linear condition for instability (4.2) is first satisfied that nonlinear eddies emerge and erode the background potential vorticity gradient in Figure 4.3. Thus, while the linear instability condition is not valid in the fully nonlinear, turbulent regime, it is nonetheless successful in predicting the initial regions of eddy stirring and subsequent emergence of mixing barriers in the adjacent regions where the background potential vorticity gradient is enhanced.

In the following sections, the statistics of the eastward and westward jets are compared and contrasted in order to isolate effects associated with strong mixing barriers.

4.4 Scaling of the mean and standard deviation

In this section we consider the relationship between the mean potential vorticity, $\langle q \rangle$, and the time-averaged streamfunction, $\overline{\psi}$, as well as the relationship between the standard deviation of potential vorticity and the time-averaged streamfunction. The angled brackets, $\langle\rangle$, imply an average over all points in time and space coincident with a contour of time-averaged streamfunction. For the numerical simulations presented here it is important to note that the time-averaged streamfunction is zonally symmetric since the simulations are run for a sufficiently long time in steady state. Therefore there is a one to one mapping between meridional distance and time-averaged streamfunction. We include, for reference, the dependence of the mean potential vorticity and the time-averaged streamfunction on meridional distance for simulations A-G in

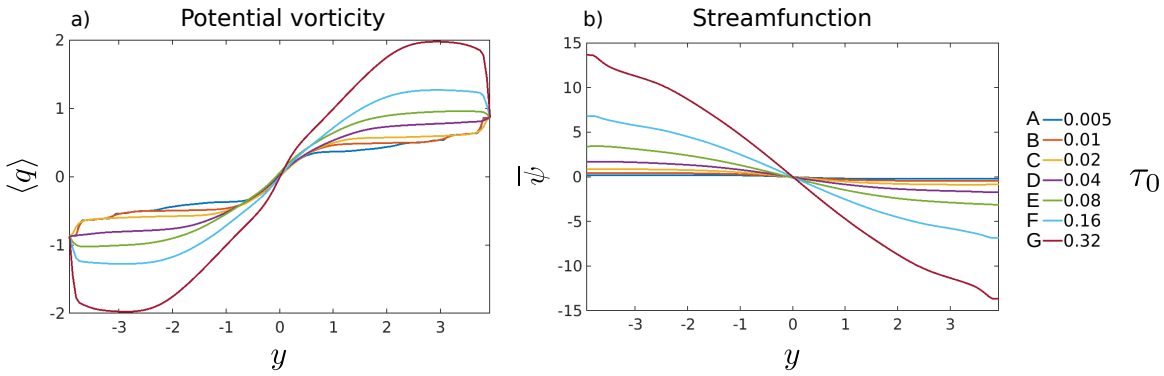


Figure 4.6: Time-averages for simulations A-G: a) time-averaged potential vorticity against meridional distance and; b) time-averaged streamfunction against meridional distance. Note, given a sufficiently long time the time-averages become zonally symmetric.

Figure 4.6.

In Figure 4.7a the $\langle q \rangle - \overline{\psi}$ relations for the forcing experiments, A-G are shown. To view the $\langle q \rangle - \overline{\psi}$ relation on the same scale, the time-averaged streamfunction has been rescaled by its average steady state boundary value $\overline{\Gamma}$. Similarly the standard deviation, SD(q), as a function of time-averaged streamfunction is presented in Figure 4.7b where

$$\mathrm{SD}(q) = \langle (q - \langle q \rangle)^2 \rangle. \tag{4.3}$$

We see a characteristic functional dependence of both the mean and the standard deviation on the time-averaged streamfunction. There is a stark contrast between these relations and those predicted by the equilibrium energy-enstrophy statistical mechanics, which predicts a linearly varying mean and a constant standard deviation, see Section 3.3.3 (e.g. Salmon et al., 1976; Jung et al., 2006). However, the mean potential vorticity dependence on streamfunction bears resemblance to the tanh-like relations ubiquitous in the generalized equilibrium Robert-Sommeria-Miller theory (Venaille and Bouchet, 2011). This implies that these statistics require more than knowledge of the energy and enstrophy to be explained. A detailed consideration for the applicability of statistical mechanics theory to forced-dissipative flow is left

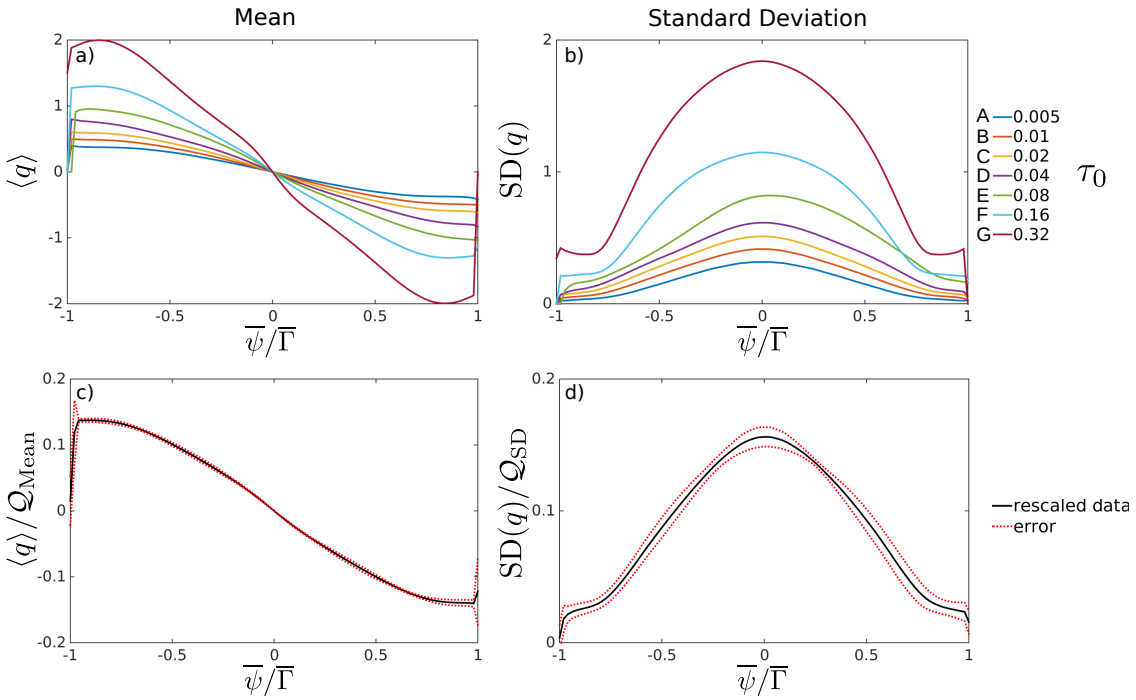


Figure 4.7: a) Mean potential vorticity against normalized time-averaged streamfunction. b) Standard deviation of potential vorticity against normalized time-averaged streamfunction. The streamfunction is normalized by it's steady state boundary value, $\overline{\Gamma}$, which is the integrated zonal momentum. Each plot shows the results from seven experiments each corresponding to a doubling of the forcing strength from the previous. The data from panels a and b are rescaled with respect to their norm to give; c) the average across all seven experiments of the rescaled mean potential vorticity as the solid-black curve; d) the average across all seven experiments of the rescaled standard deviation of the potential vorticity as the solid-black curve. In both c and d the dashed-red curves give the standard error.

for future studies. Here we highlight the generic nature of these statistics and their behaviour over a wide range of forcing and dissipation.

Figure 4.7c and d shows that the mean and standard deviation of potential vorticity display a self-similar dependence on the normalized streamfunction when both the $\langle q \rangle - \overline{\psi}$ relation and the $\text{SD}(q) - \overline{\psi}$ relation are rescaled by their l^2-norm. The red dashed lines in Figure 4.7c and d gives the error in this self-similarity which is found to be small. Interestingly it is possible to relate the normalization factors for $\langle q \rangle$ and $\text{SD}(q)$ to the enstrophy of the flow. Assuming zonal symmetry in the time-mean, we

can define a mean flow enstrophy as

$$C_2[\langle q \rangle] = L_x \int \mathrm{d}y \, \langle q \rangle^2 \tag{4.4}$$

and an eddy enstrophy as

$$C_2[q'] = L_x \int \mathrm{d}y \, \left[\langle q^2 \rangle - \langle q \rangle^2 \right] \tag{4.5}$$

where $q' = q - \langle q \rangle^2$, $C_2[q] = C_2[\langle q \rangle] + C_2[q']$ and L_x is the zonal extent of the domain and enters through integrating over the zonal direction.

For the mean, the normalization factor is given as the l^2-norm of the curves in Figures 4.7a and b, which is written as

$$\mathcal{Q}_{\mathrm{Mean}} = \sqrt{\int \mathrm{d}y \, \langle q \rangle^2} = \sqrt{\frac{C_2[\langle q \rangle]}{L_x}}, \tag{4.6}$$

while for the standard deviation the normalization factor is similarly given as

$$\mathcal{Q}_{\mathrm{SD}} = \sqrt{\int \mathrm{d}y \, \left[\langle q^2 \rangle - \langle q \rangle^2 \right]} = \sqrt{\frac{C_2[q']}{L_x}}. \tag{4.7}$$

We find that the $\langle q \rangle - \overline{\psi}$ relation and the $\mathrm{SD}(q) - \overline{\psi}$ relations are self-similar over a wide range of eastward forcing strengths (Figure 4.7), but not when the magnitude of the dissipation is varied, or the sign of the forcing is changed, causing the jet to flow in the opposite direction (Figure 4.8). The perturbed dissipation experiments, if anything, show limited response in the mean while there is a scaled, although less clean, response in the standard deviation of the potential vorticity. Although there is still a partial self-similarity the collapse is found to be less clean. We suggest that this difference is related to the fact that scale dependence of the forcing is unchanged when its strength is varied, it always acts at the first zonal wavenumber. On the other

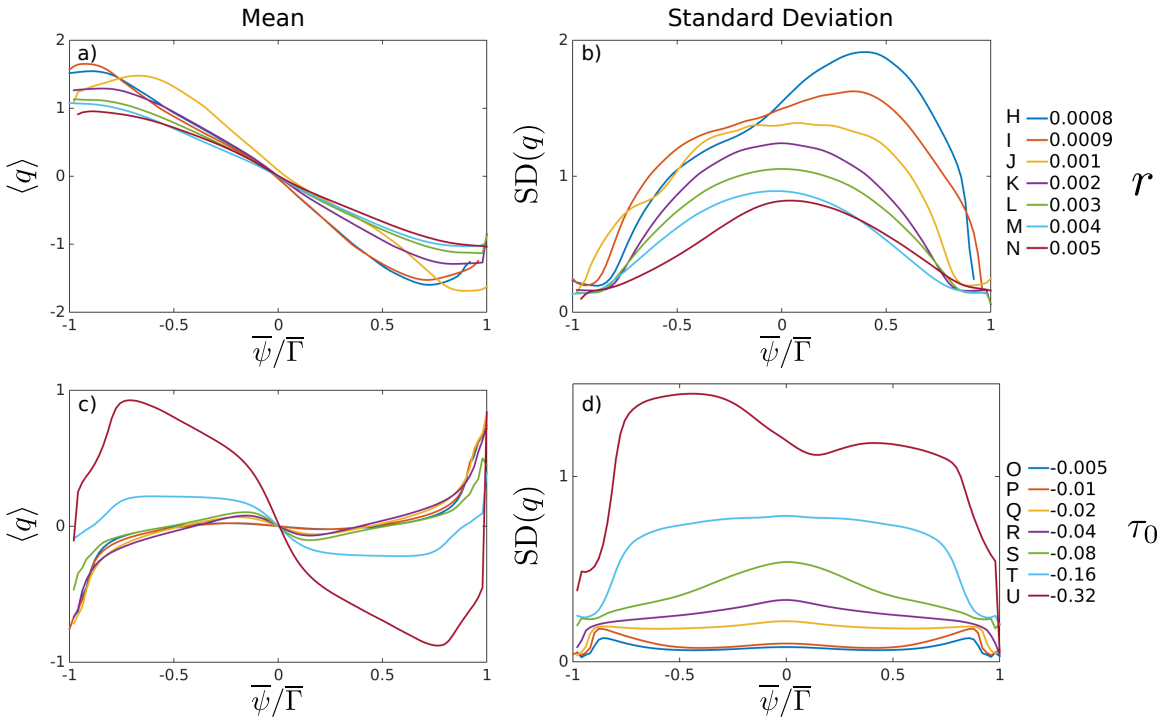


Figure 4.8: The mean (a and c) and standard deviation (b and d) as functions of time-averaged streamfunction are shown for the dissipation (a and b) and reversed forcing (c and d) experiments.

hand, the linear drag term interacts strongly with the flow over a different range of scales as the linear drag coefficient is varied. Investigating this effect would require additional experiments in which the width of the wind stress profile is varied as well as the structure of the wind forcing; this is not explored further here.

An interesting feature of the experiments in which the dissipation is varied (Figure 4.8a and b) is the asymmetry in the SD$(q) - \overline{\psi}$ relation about $\overline{\psi} = 0$ for the low dissipation simulations, in particular H and I. We can see a clue to the resolution of this puzzle in Figure 4.9 which shows a snapshot of steady state potential vorticity for simulation H, the simulation with the lowest linear drag coefficient. We find that a large coherent eddy has emerged in bottom half of the domain. A similar eddy can also be seen in simulations I and J, see Figure 4.4; in the case of J the eddy appears in the upper part of the domain. It is unclear why this eddy forms when the friction is decreased.

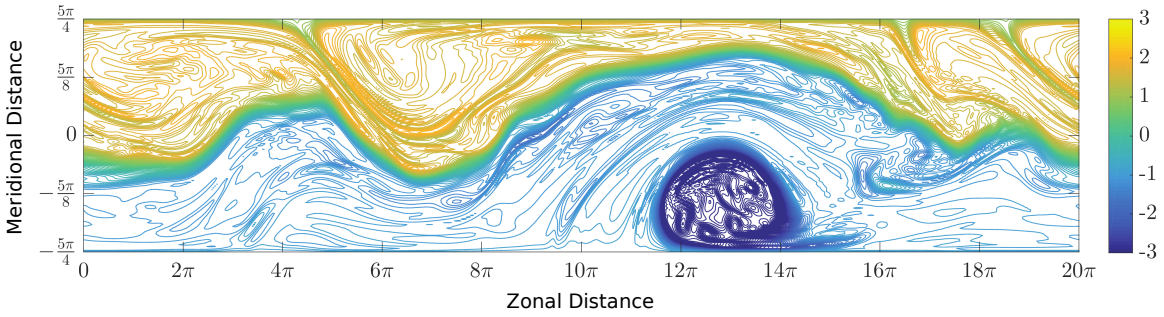


Figure 4.9: Snapshot of steady state potential vorticity for simulation H, $r = 0.0008$. A large eddy with very negative potential vorticity is persistent in the bottom half of the domain.

When the direction of the forcing is reversed, the mean and standard deviation dependence on the time-averaged streamfunction is drastically altered and again there is some self-similarity, although this is much less clear than for the eastward jets. The main physical difference in the reversed forcing experiments is the loss of the clear mixing barrier formed through Rossby wave dynamics. When the forcing is reversed the instabilities are strong enough to ensure that a persistent Rossby wave is broken up and this changes the nature of the flow in a significant manner.

In summary, it is found for eastward forcing that the mean and standard deviation of the potential vorticity scales with forcing and that the dependence of the mean and standard deviation on the time-averaged streamfunction is self-similar over a large range of forcing strengths. The $\langle q \rangle - \overline{\psi}$ relation is self-similar when the streamfunction is scaled by the total through channel momentum and the potential vorticity is scaled by the square root of the mean field enstrophy. The $\mathrm{SD}(q) - \overline{\psi}$ relation is self-similar when the streamfunction is scaled by the total through channel momentum and the potential vorticity is scaled by the square root of the eddy enstrophy. Thus, for this experimental set-up it is possible to empirically deduce the mean and standard deviation given that we know the enstrophy behaviour of the flow. It is inconclusive whether the mean scales with dissipation but there is clear response in the standard deviation to changed dissipation, although the self-similarity is not as clean. We see

a changed behaviour in the reversed forcing experiment and leave a full description of this flow regime for future work. In the following sections we concentrate on describing the statistical phenomenology for the forcing experiments, A-G, referring to the other experiments when pertinent to the argument.

4.5 Skewness, kurtosis and jet meandering

Gaussian statistics can be characterized using only knowledge of the mean and standard deviation. When the statistics are non-Gaussian it becomes important to consider higher order statistics. We will here consider the 3^{rd} and 4^{th} order centred and standardized moments, the skewness and kurtosis[1], given respectively as

$$S = \frac{\langle (q - \langle q \rangle)^3 \rangle}{\mathrm{SD}(q)^3} \tag{4.8}$$

and

$$K = \frac{\langle (q - \langle q \rangle)^4 \rangle}{\mathrm{SD}(q)^4}. \tag{4.9}$$

The skewness is a measure of how asymmetric or lopsided a statistical distribution is about its mean while the kurtosis is a measure of 'peakiness'. A high kurtosis implies a distribution which has a sharp peak with fat tails, a low kurtosis implies small tails but a flat peak[2]. Figure 4.1 illustrates these moments schematically. Plots of the skewness and kurtosis against normalized time-averaged streamfunction are given in Figure 4.10a and b, for the simulations A-G. Contrasted with the scaling behaviour of the mean and standard deviation, for central values of the normalized time-averaged streamfunction, $-0.4 < \overline{\psi}/\overline{\Gamma} < 0.4$, the skewness and kurtosis display invariance with respect to changed forcing. Not only is the skewness and kurtosis in

[1]Note that in some literature the term kurtosis is used to mean what is described here as the excess kurtosis, $K - 3$.

[2]This picture relies on having an uni-modal distribution with one peak.

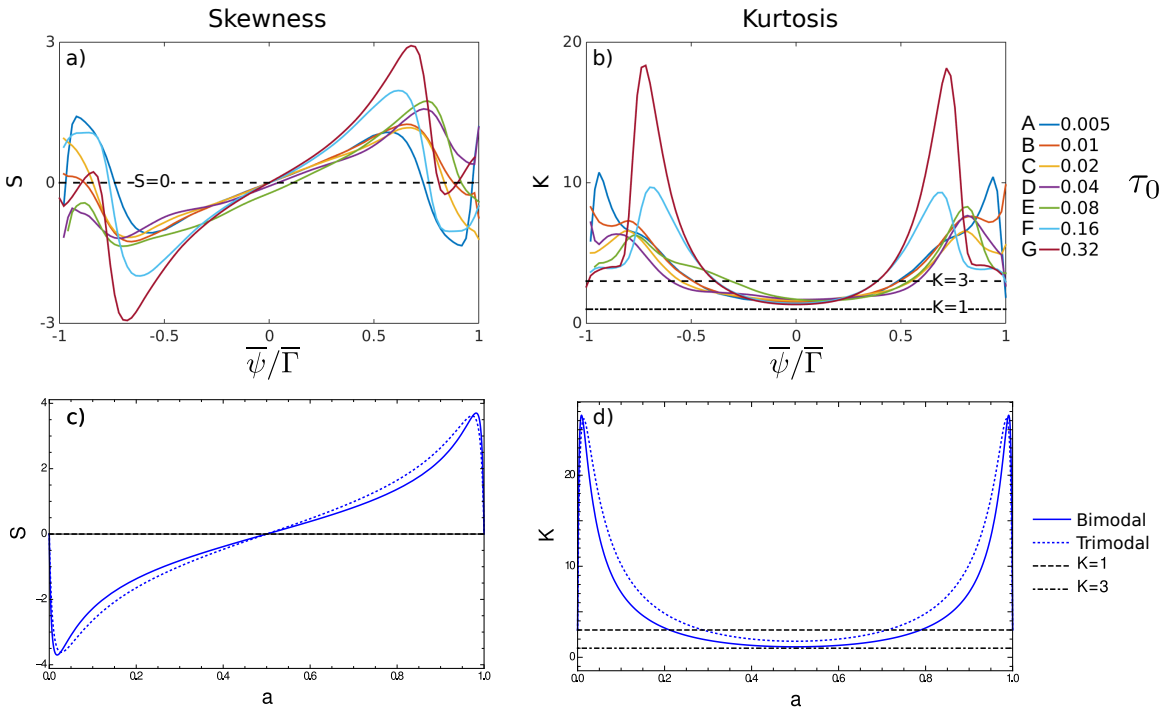


Figure 4.10: a) skewness of potential vorticity as a function of time-averaged streamfunction. The line $S = 0$ is shown and represents statistics which is symmetric about its mean. b) kurtosis of potential vorticity as a function of time-averaged streamfunction. The lines $K = 3$ and $K = 1$ are shown. A kurtosis of 3 corresponds to a Gaussian distribution, a kurtosis greater than this is referred to as leptokurtic and while a kurtosis of greater than three is referred to as platykurtic. A kurtosis of 1 is the smallest possible kurtosis and corresponds to an unbiased coin toss. In c and d the skewness and the kurtosis are given respectively from the bimodal distribution, (4.10), and the trimodal distribution, (4.12). The parameters used are $d = 10$ and $c = 0.6$.

the jet core behaviour fixed with respect to forcing strength, but this behaviour can interpreted employing the results of Hughes et al. (2010) who show that a change in sign of skewness coincident with low kurtosis values is indicative of a mixing barrier. This is shown both in the relative vorticity anomaly data inferred from altimeter observations as well as in a two-layer quasi-geostrophic model in a doubly periodic channel. In this study, the same skewness/kurtosis behaviour is seen in a barotropic model in a singly-periodic channel.

Central to the argument of Hughes et al. (2010) is the idea that a jet which acts as a mixing barrier leads to a step in potential vorticity with homogenized values

of potential vorticity on either side of the barrier. This potential vorticity staircase paradigm for jets is reviewed in Dritschel and McIntyre (2008) and paints a picture (along with the references therein) in which mixing or homogenization of potential vorticity either side of a jump in the value of potential vorticity acts to reinforce the mixing barrier and to amplify the jump in potential vorticity. This feedback leads us to a picture where an observer standing at the centre of the jet, defined by either $\overline{\psi} = 0$ or $y = 0$, would see the value of potential vorticity flip from negative to positive as the mixing barrier meanders past due to non-linearities or Rossby waves.

With the potential vorticity step picture in mind, we can explain the skewness/kurtosis behaviour shown in Figure 4.10. A kurtosis of $K = 1$ is described as sub-Gaussian. Kurtoses of greater than three are described as leptokurtic and kurtoses of less than three are described as platykurtic. In Figure 4.10, the skewness plot is marked with the line $S = 0$. The kurtosis plot is marked with the $K = 3$ line corresponding to a Gaussian distribution and the $K = 1$ line which corresponds to an unbiased coin toss or the unbiased Bernoulli process, a stochastic process where the random variable can take one of two values. We see a distinct change in the value of skewness for a time-averaged streamfunction of around zero. The skewness behaviour is consistent across all forcing strengths for a range of $-0.4 < \overline{\psi}/\overline{\Gamma} < 0.4$, which we refer to as the core of the jet. For larger values of normalized time-averaged streamfunction the skewness becomes more varied across the different simulations. Coincident with the described behaviour of skewness we see similar consistency in the kurtosis for the range $-0.4 < \overline{\psi}/\overline{\Gamma} < 0.4$. The kurtosis in the core of the jet is predominately platykurtic and we also see the kurtosis becoming close to sub-Gaussian. This skewness/kurtosis behaviour is consistent with the findings of Hughes et al. (2010) and can be interpreted as due to a flip in the value of potential vorticity as the mixing barrier meanders, through the action of passing Rossby waves.

In fact, we can go further than the above discussion and consider a simple ana-

lytical model for the skewness and kurtosis used by Hughes et al. (2010). This model describes the statistics as Gaussian either side of a step in the value of potential vorticity, with the meandering of the jet represented as a Bernoulli process. Thus, the *centred* probability distribution for the potential vorticity is given by the sum of two normal distributions scaled by factors, a and $1-a$ and separated by d times their standard deviation, σ. The parameter, a, is a decreasing function of meridional distance to represent the decreasing probability of observing positive potential vorticity as the jet is crossed. The probability distribution is

$$\rho(q) = (1-a)\mathcal{N}(q,\, a\sigma d,\, \sigma) + a\mathcal{N}(q,\, (a-1)\sigma d,\, \sigma) \tag{4.10}$$

where

$$N(q, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q-\mu)^2}{\sigma^2}\right), \tag{4.11}$$

is the normal distribution as a function of potential vorticity, q, which is treated as a random variable. The distribution is parameterized by a mean, μ, and standard deviation, σ. By making the mean of each of the Gaussian peaks depend on a we ensure the total distribution has zero mean, giving us the centred statistics we need to calculate the skewness and kurtosis. The skewness and kurtosis corresponding to this distribution is shown in Figure 4.10c and d respectively. These analytically derived curves bear a striking resemblance to the observed behaviour as a function of streamfunction shown in Figure 4.10a and b (and also for the corresponding curves in Figure 4.11 for the dissipation experiment only). The observed behaviour of skewness/kurtosis is more variable near the boundary, suggesting that the simple statistical model presented here is most appropriate near the core of the jet.

This simple model is able to qualitatively reproduce the skewness/kurtosis behaviour extending away from the centre of the jet. We can intuitively explain the high peaks in kurtosis at large values of streamfunction; as one of peaks gets very

small the kurtosis it appears as a very large skewed tail of the larger peak leading to a high value of kurtosis. However, as the height of the smaller peak goes to zero, the kurtosis must go to three; that is, the kurtosis of a normal distribution. Due to the close similarity between the analytical curves and the model diagnosed curves, it is tempting to forget that the former is plotted as a function of normalized time-averaged streamfunction, $\overline{\psi}/\overline{\Gamma}$, while the latter is a function of the weighting factor, a. Thus the resemblance suggests a close relationship between a and $\overline{\psi}/\overline{\Gamma}$, both monotonic functions of meridional distance, although the establishment of a quantitative relationship remains an open problem.

One matter of concern lies in the fact that when we plot the distribution of potential vorticity for a given time-averaged streamfunction the statistics is trimodal, Figure 4.12, unlike the bimodal distribution, (4.10). To take account of this, the calculation of Hughes et al. (2010) is modified by a central Gaussian peak of height weight, c, giving a probability distribution of

$$\begin{aligned} \rho(q) &= \frac{1}{1+c}\Big[(1-a)\mathcal{N}(q,\, a\sigma d,\, \sigma) \\ &\quad + c\mathcal{N}(q,\, 0,\, \sigma) \\ &\quad + a\mathcal{N}(q,\, (a-1)\sigma d,\, \sigma)\Big]. \end{aligned} \tag{4.12}$$

The results of this calculation is also shown in Figure 4.10. The trimodal nature of the statistics observed in some of the experiments will be discussed further in Sections 4.6 and 4.7. However it is clear from Figure 4.10 that the addition of a central peak does not change the qualitative story.

In a barotropic channel, we have observed a change in sign of skewness across a jet, this, combined with a low kurtosis, is indicative of a mixing barrier. A highly simplified model describing the statistics of the jet meandering as a Bernoulli distribution with superimposed small-scale Gaussian noise gives surprisingly good qualitative

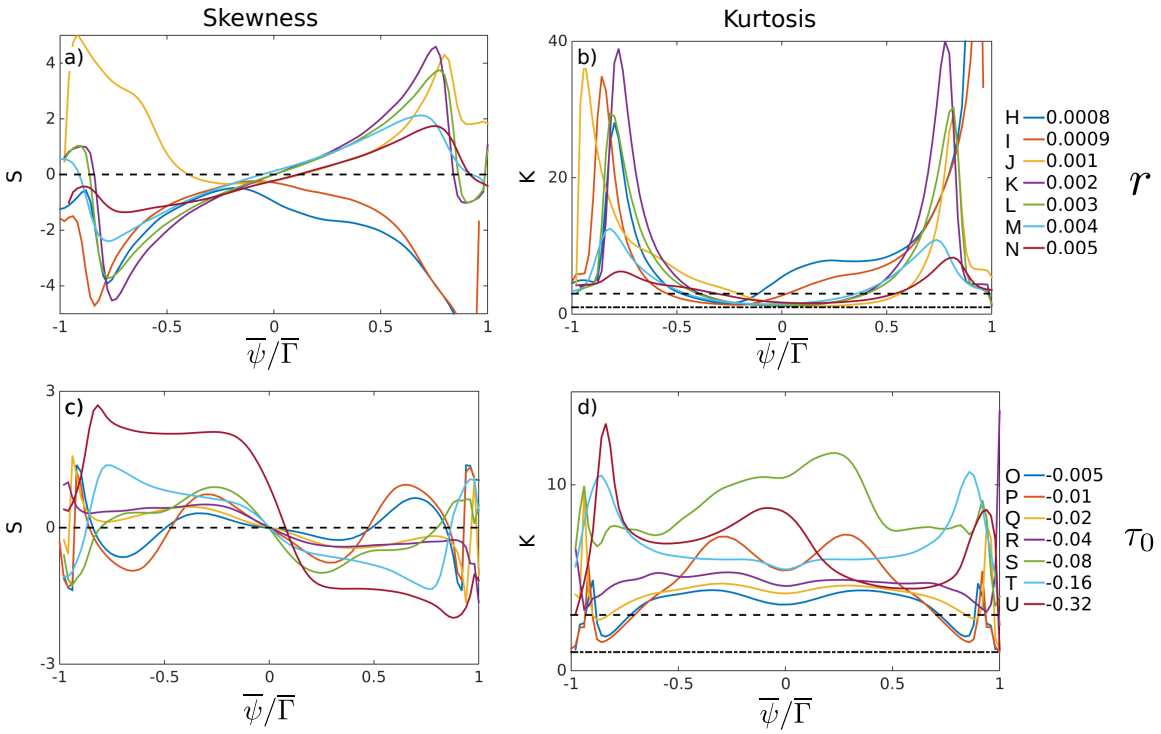


Figure 4.11: The skewness (a and c) and kurtosis (b and d) as functions of time-averaged streamfunction are shown for the dissipation (a and b) and reversed forcing (c and d) experiments. Again the $S = 0$ line is shown for the skewness and the $K = 3$ and $K = 1$ for the kurtosis.

agreement with the observed skewness and kurtosis across the jet over a wide range of forcing and dissipation.

Figure 4.11 shows how this behaviour changes in the dissipation and reversed forcing experiments. For the dissipation experiments, Figure 4.11a and b, we see a similar picture to the forcing experiments for all but the two lowest dissipation parameters. These outliers are linked to the clear asymmetry in the standard deviation shown in Figure 4.8b. For the other dissipation parameters we see a crossing of the $S = 0$ line coincident with platykurtic kurtoses close to $K = 1$. The real difference is seen in the reversed jet experiment (Figure 4.11c and d). Here the skewness behaviour is in the opposite sense, now going from positive to negative, and the values differ greatly between the different (reversed) forcing strengths. In the core of the jet the kurtosis is leptokurtic for all simulations and does not display the near sub-Gaussian behaviour

in the other two experiments.

The drastic change of the skewness/kurtosis behaviour can be understood when we realize that in reversing the direction of forcing, resulting in a east to west rather than a west to east jet, we have changed the jet dynamics, as discussed in Section 4.3, causing the mixing barrier to be broken up. This allows the potential vorticity to be mixed in the north-south direction, resulting in a jet without a mixing barrier, destroying the coin-toss like picture and the corresponding signal in the skewness/kurtosis behaviour. When the mixing barrier is present the skewness/kurtosis behaviour can be explained by a mixed distribution; this will be discussed further in Section 4.7. We have shown here that in a barotropic re-entrant channel that the presence of a mixing barrier results in a characteristic skewness/kurtosis across jet behaviour. When the mixing barrier is removed, the skewness/kurtosis behaviour does not persist.

4.6 Streamfunction-following statistics

In the previous section we saw that the skewness/kurtosis behaviour can be partially explained through a simple statistical description in the presence of a mixing barrier and that the observed skewness/kurtosis behaviour is indeed characteristic of a mixing barrier. In this section, we consider the dependence of the statistics on the coordinate system in which the statistics are calculated. We consider three frames of reference, leading to three methods of time-averaging:

- Eulerian time-averaging, in which the statistics are calculated along the zonally-symmetric, time-mean streamfunction contours.

- Galilean transformed time-averaging, in which the statistics are calculated along the non-zonally symmetric time-mean streamfunction contours in a frame of reference moving at the phase speed of the dominant Rossby wave. Further details of this methodology as well Rossby wave spectra, identifying waves other

than the dominant wave, are presented in Appendix B,

- Streamfunction following time-averaging, in which the statistics are calculated along the instantaneous streamfunction contours.

In the interests of brevity, we present results only for the experiments with the changed forcing strength, A-G.

We start by considering in more detail the statistics of the flow for a particular simulation. Figure 4.12 shows the various ways to view histograms of the potential vorticity for simulation D from the forcing experiment. At the top of the figure the joint histogram for potential vorticity and streamfunction is shown for all zonal distances and times on the $y = 0$ (or approximately the $\overline{\psi} = 0$ line). The histogram is trimodal, looking like a mountain range with three peaks running in a fairly straight $q - \psi$ line. We see here three distinct statistical populations in q and ψ. Below in Figure 4.12 are histograms of potential vorticity using various ways of taking the statistics. On the left we have shown the histogram of potential vorticity for all zonal distances and times on the $\overline{\psi} = 0$ line. This is equivalent to the joint histogram above where we have marginalized out the streamfunction. We see that although the marginalized histogram has three peaks the joint histogram tells us that each peak corresponds to different values of instantaneous streamfunction.

In what follows, it will be shown that the two side peaks correspond to the regions of homogenized potential vorticity either side of the potential vorticity step. This gives rise to the skewness/kurtosis behaviour considered in Section 4.5. In the middle plots, we consider how moving to a frame of reference travelling at the wave-speed of the most dominant Rossby wave may affect the statistics. In this way we hope to remove much of the meandering of the jet that is due to the dominant Rossby wave passing through the channel. This yields the Galilean transformed time-averaged streamfunction shown in Figure 4.12 (*bottom-middle*). Despite changing the relative size of the three peaks discussed previously, this method of taking the statistics is

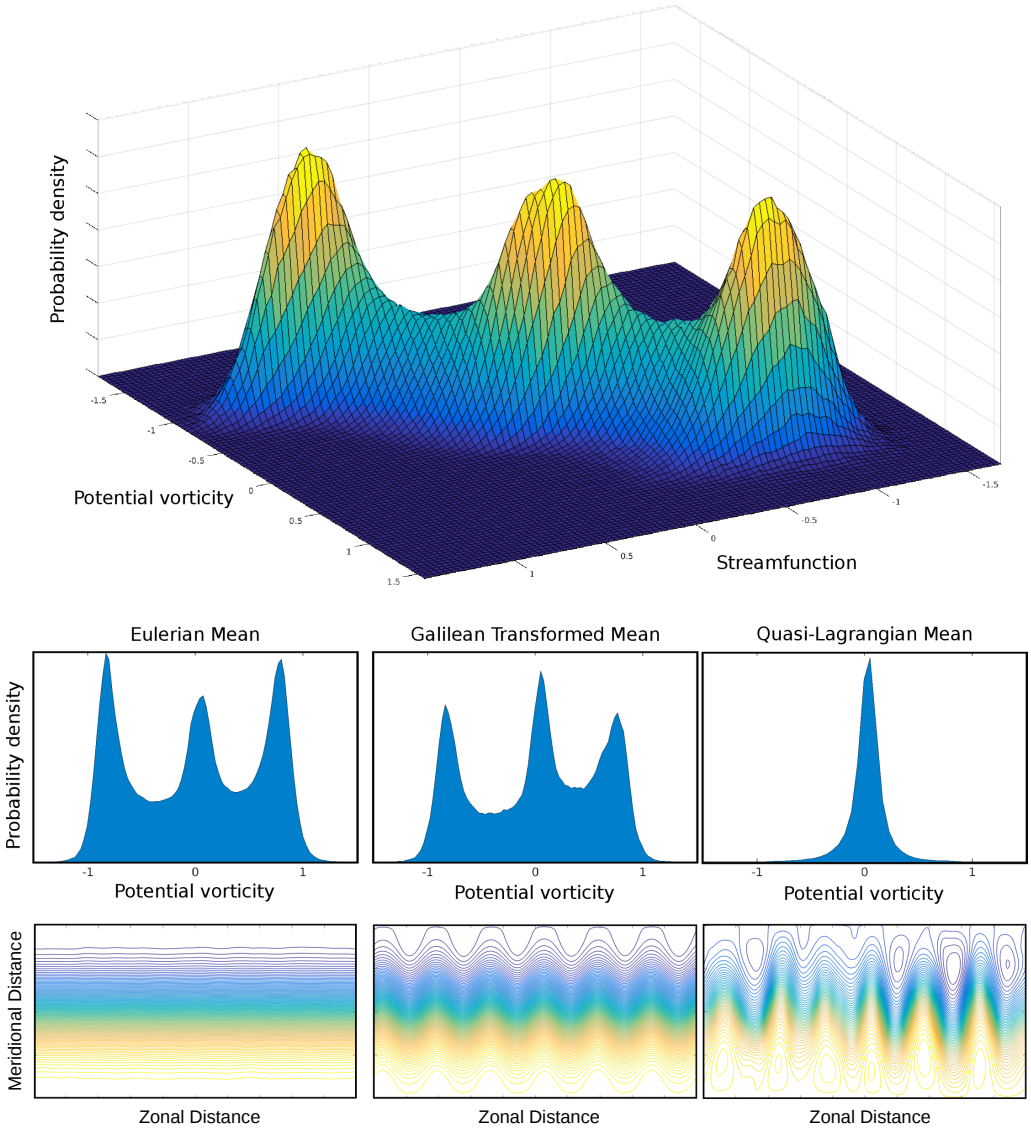


Figure 4.12: This figure shows different ways to look at the statistics in the middle of the jet for forcing experiment D. *Top* - joint histogram of potential vorticity and streamfunction at the central meridional distance (close to the $\overline{\psi} = 0$ line). *Middle row* - three histograms of potential vorticity with the statistics sampled in different ways. *Bottom row* - streamfunction in various frames of reference are plotted illustrating the ways the statistics were sampled for the middle histograms. *Left column* - statistics taken from positions defined by the time-averaged streamfunction, $\overline{\psi} = 0$ contour, these positions are fixed in time. *Middle column* - statistics taken from positions defined by the Galilean transformed time-averaged streamfunction, $\overline{\psi_{GT}} = 0$ contour, these positions are also fixed in time. *Right column* - statistics taken from positions defined by the instantaneous streamfunction, $\psi = 0$ contour, these positions are allowed to change with time. The scale for the distribution above is omitted as it can be normalized away and we are only interested in the shape of the distribution at present.

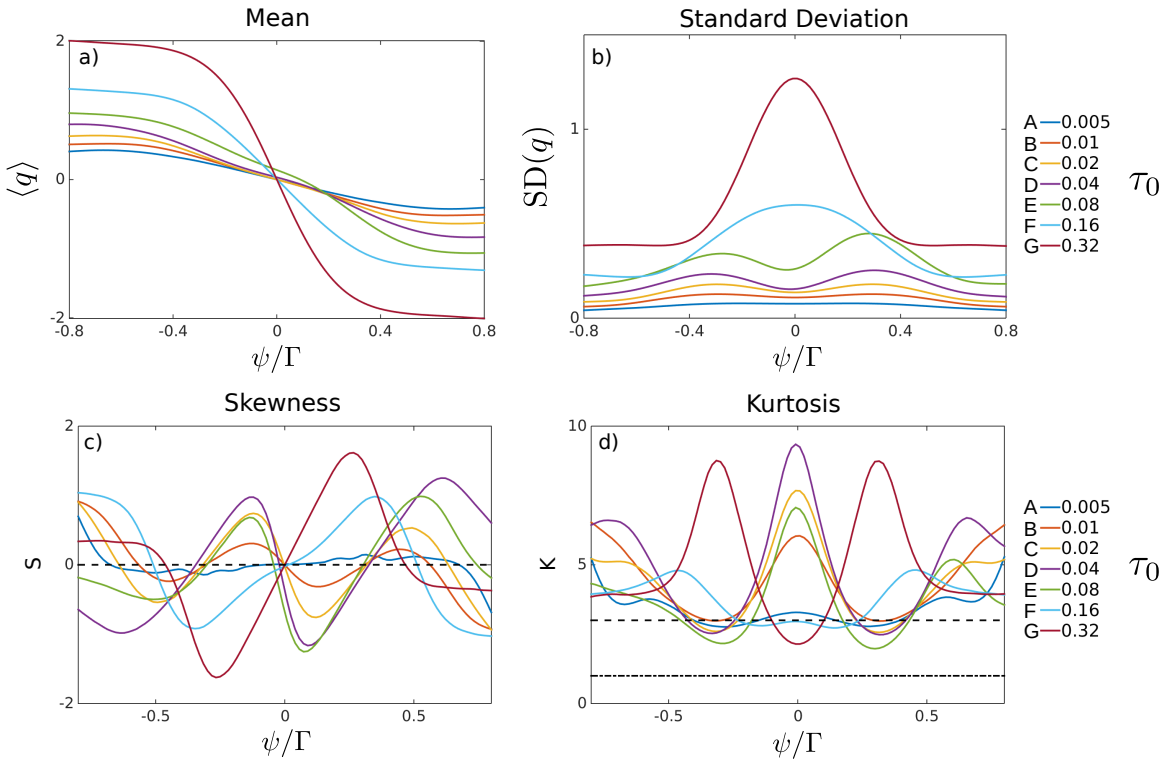


Figure 4.13: a) the mean, b) standard deviation, c) skewness and d) kurtosis as functions of instantaneous streamfunction are show. Here the statistics has been sampled along contours of streamfunction for all times.

insufficient to remove the tri-modality of the histogram. If however, we use the instantaneous streamfunction $\psi = 0$ as the spatial ordinate for taking the statistics we completely remove the two side peaks in the distribution giving rise to an unimodal histogram. Thus by explicitly moving with the meandering jet we have drastically altered, and simplified, the structure of the potential vorticity distribution. This also proves that the the two side peaks, apparent when the statistics is viewed from an Eulerian perspective, are due to the large scale meandering of the mixing barrier and are distributions of the regions of homogenized potential vorticity either side. This leaves only the potential vorticity statistics within the strong potential vorticity gradient itself as can be seen in Figure 4.12.

Figure 4.13 shows the first four moments considered in previous sections now as a function of the normalized *instantaneous* streamfunction. The behaviour of the

mean is qualitatively unchanged from that observed in Section 4.4, while the standard deviation displays major qualitative differences. For the lower forcing strengths A-E the standard deviation is suppressed around $\psi = 0$ while for the two highest forcing strengths the central standard deviation is enhanced. The variation between simulations in the skewness and kurtosis is very large and it is hard to discern a consistent behaviour. However it clear that characteristic behaviour indicative of a mixing barrier is now lost and the kurtosis is on the whole leptokurtic and does not show the low (close to $K = 1$) kurtosis behaviour when the statistics was taken on contours of time-averaged streamfunction. The change of sign of skewness is no longer coincident with low kurtosis values. Thus, we have shown that this behaviour disappears when the mixing barrier is removed either through changing the physics or by moving into a frame of reference of the instantaneous streamfunction.

At this point it is worth considering how these statistics, given in Figure 4.13, correspond to the prediction of equilibrium statistical mechanics theory. An equilibrium, energy-enstrophy, theory (Salmon et al., 1976; Jung et al., 2006) predicts that the mean will be linear with respect to streamfunction, the standard deviation will be constant, the skewness will have value, $S = 0$, and the kurtosis will have value, $K = 3$. We see that these four criteria are approximately met in the centre of the jet for the lowest strength of forcing. This suggests agreement with the results of the equilibrium theory when forcing is weak *and* the meandering of the mixing barrier has been taken into account. This chimes with the study of Jung et al. (2006) in which they moved into the frame of reference of their large scale Rossby wave to observe their statistics. We suggest that the reason this procedure does not work here is the absence of any significant scale separation in contrast to the experiment of Jung et al. (2006). This could be due to various factors including: strength of dissipation; length scale of forcing; and strength of the β-term.

The joint histograms agree with the interpretation of separate statistical popula-

tions either side of zero streamfunction and is consistent with presence of a mixing barrier. Here we considered one example of the joint q-ψ statistics but will consider the histograms for other simulations further in Section 4.7. Although containing much Lagrangian information the streamfunction-following frame of reference is not equivalent to following a Lagrangian trajectory for a time-varying streamlines, we will refer to such a frame of reference as *quasi-Lagrangian*. We have shown that performing the statistical analysis in a quasi-Lagrangian frame of reference removes the multimodal nature of the distribution simplifying the statistics but at the cost of a more complicated time-averaging. In addition, the simulation with lowest forcing strength gives an approximate agreement with Jung et al. (2006) in both the quasi-Lagrangian frame and the reversed forcing statistics. It could be argued that as the strength of forcing is reduced the statistics tends to the (energy-enstrophy) equilibrium statistical mechanics prediction, but only when the mixing barrier is removed.

4.7 Discussion

In this discussion we will draw together the results from the previous sections by considering: joint distributions of potential vorticity and streamfunction; the concept of mixed distributions and their importance to describing mixing barriers[3]; implications for stochastic parameterizations; and the application to the Southern Ocean. In this way we are able to highlight some key implications of the results of this study.

4.7.1 Joint probability distributions

To highlight the statistical nature of mixing barriers it is fruitful to consider the $q-\psi$ joint histograms for all the forcing strengths in simulations A-G, given in Figure

[3]Due to an unfortunate combination of terminology from different fields the prefix 'mix' here should be understood in different ways; in the former case 'mixed' refers to a mix of continuous and discrete random variables, in the latter case 'mixing' refers to turbulent mixing.

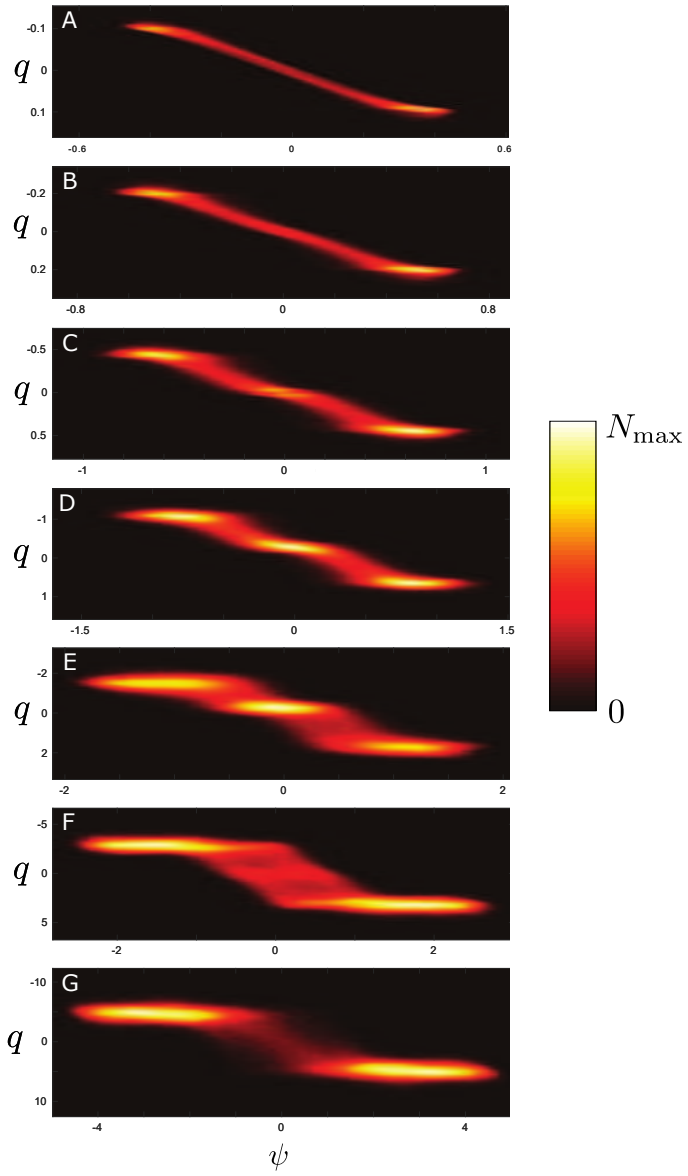


Figure 4.14: Plots of the joint histograms of potential vorticity and streamfunction for the forcing experiment, simulations A-G. Colour scale gives the number of counts in each of the 100×100 bins. *Brighter - more counts. Black - zero counts. White - maximum counts.*

4.14. As with the example given in Figure 4.12 we have plotted the joint counts of potential vorticity and streamfunction for 100×100 bins, with the width set by the range of the data. As mentioned, the distributions are a mixture of bimodal and trimodal; A and G are bimodal; B and F have a weak central peak; C, D and E are trimodal. As described previously we see the clear separation of the statistical

populations consistent with the picture of a meandering potential vorticity step, but further to this we see interesting features of the statistics.

Firstly the question is raised of what causes the existence of the third peak seen in B-F, centered at 0. From examining the details of the flow evolution it appears that this central peak coincides with the emergence of a spatially narrow region in the core of the jet where the potential vorticity gradient is weakened coming to zero (or even a weakly negative value). This narrow homogenized region appears intermittently within the otherwise strong potential vorticity gradient associated with the channel Rossby wave. This can be seen in the snapshot from simulation D, Figure 4.2. While difficult to observe in noisy snapshots of potential vorticity, this becomes apparent when the potential vorticity is plotted against streamfunction as the streamfunction's latitudinal gradient (or zonal velocity) has its largest magnitudes near the centre of the jet, see Figure 4.12, D. Secondly we notice that the areas of high bin count are elongated in the streamfunction direction, giving the same value of mean potential vorticity for many values of streamfunction: apparent in the data from simulations C-G. This quite clearly demonstrates the potential vorticity homogenization paradigm discussed in, for example, Esler (2008a,b). Indeed, the potential vorticity step with adjacent regions of homogenization is made very apparent when the flow is viewed as a potential vorticity/streamfunction joint distribution.

4.7.2 Mixed distributions

We have shown that the statistics of potential vorticity in the presence of a mixing barrier are described by mixed distributions, a sum of distributions of random variables. Mixed distributions are distributions corresponding to random variables created by a random choice of other random variables. In other words, they are a convolution of discrete and continuous random variables. Here, the discrete random variable represents the regions separated by the mixing barrier and the continuous

random variable represents the statistics of the homogenized regions. The full Eulerian statistics is then given by randomly selecting which side of the mixing barrier we are standing through a Bernoulli trial. The simple models (4.10) and (4.12) are examples of mixed distributions. This opens up the possibility of using Gaussian Mixture Modeling methods to analyse the statistics of meandering mixing barriers. These methods have been used in atmospheric dynamics (Hannachi, 2007) to consider weather regimes and also, more recently, in oceanography (Maze et al., 2017) to classify hydrographic profiles.

An important consequence of the mixed nature of these distributions is that they are difficult to reconstruct from a maximization of information entropy given knowledge of a modest number of moments. If we assume that we know that the first N moments

$$\langle x^n \rangle = \int \mathrm{d}x x^n \rho(x) \tag{4.13}$$

of a probability distribution, ρ, then maximization of information entropy

$$S[\rho] = -\int \mathrm{d}x \rho \ln \rho, \tag{4.14}$$

gives a probability distribution of the form

$$\rho = \frac{1}{\mathcal{Z}} \exp\left(\sum_{n=1}^{N} \alpha_n x^n\right), \tag{4.15}$$

where the α_n are the Lagrange multipliers used in the maximization of entropy (e.g. Mead and Papanicolaou, 1984).

Mixed distributions such as (4.10) and (4.12) cannot be inferred in this way through only knowledge of the measured moments. The maximum entropy distribution, (4.15), can be used to approximate multimodal distributions such as (4.10) and (4.12) but this requires N to be very large (example calculations suggest N needs

to be as high as 100), considerably beyond what is available from data. Essentially this comes from the difficulty of representing a bimodal distribution by the exponential of a Taylor expansion: a very large number of terms would need to be kept to capture two well separated peaks. The maximum entropy method of approximating probability distribution functions was successfully used in Porta Mana and Zanna (2014) to fit the diagnosed distribution of a stochastic parameterization where the distributions considered were unimodal; this method of fitting probability distribution functions would be inappropriate for use in the case considered in this study.

4.7.3 Stochastic parameterization

The results in this study have important consequences for the way we approach the stochastic parameterization of ocean mesoscale turbulence. If we parameterize the Eulerian eddy forcing tendency, then it is essential that we do this in a way that will lead to a mixed probability distribution for potential vorticity in the presence of a strong mixing barrier. It is therefore prudent to both take into account higher order cumulants as well as spatial correlations as these are required to describe the structure of jets. This is the approach followed in stochastic structural stability theory (e.g. Farrell and Ioannou, 2003) and in the cumulant expansion method (e.g. Marston et al., 2008); by solving prognostic equations for statistical cumulants of the flow it is possible represent eddy-mean flow interactions which respect the steady state structure and dynamics of turbulent jets.

An alternative interpretation of the results presented in this study is that a stochastic parameterization should be implemented in a Lagrangian sense, allowing the noise to be simpler in structure: monomodal instead of the multimodal statistics seen in a Eulerian frame. Recently parameterizations have been developed which modify the Lagrangian tendencies rather than the Eulerian tendencies (Holm and Nadiga, 2003; Holm and Wingate, 2005; Porta Mana and Zanna, 2014; Zanna et al.,

2017); Porta Mana and Zanna (2014) do this by *stochastically* modifying the Lagrangian tendency of potential vorticity in a quasi-geostrophic model. We can see this by considering the potential vorticity equation

$$\frac{Dq}{Dt} = D + F + \eta, \tag{4.16}$$

where D, F and η are the dissipation, forcing and stochastic noise respectively. In this formulation the noise can only influence the Eulerian tendency of potential vorticity. However, the parameterization introduced by Porta Mana and Zanna (2014) can be written as

$$\frac{Dq}{Dt} = D + F + \kappa\nabla^2\left(\frac{Dq}{Dt}\right) \tag{4.17}$$

where κ is now a stochastic noise with a negative mean value. By rewriting as

$$\left(1 - \kappa\nabla^2\right)\frac{Dq}{Dt} = D + F, \tag{4.18}$$

we can see that the additive noise in (4.16) cannot be distinguished from a perturbation of the Eulerian tendency only, whilst the multiplicative operator in (4.18) can. This makes parameterizations of this form an advantageous framework for a stochastic treatment as the potential vorticity distributions become simpler when viewed in a Lagrangian frame of reference, see Section 4.6.

4.8 Concluding remarks

The relationships between streamfunction and the various statistics we have diagnosed in this study cannot be explained by equilibrium energy-enstrophy statistical mechanics theories, contrary to the result of Jung et al. (2006) who found good agreement with a barotropic rotating annulus experiment. Whether more general versions of the theory can be used to explain the generic statistical characteristics of this flow

will require a deeper understanding of the response of the statistics to forcing and dissipation. In this study we explored the phenomenology of these statistics over a wide range of forcing and dissipation strengths. We have shown that the mean and standard deviation of the potential vorticity as a function of time-averaged streamfunction exhibits self-similar behaviour for barotropic flow across a wide magnitude range of eastward wind stress. The self-similarity is also observed when the dissipation is perturbed and when the direction of wind stress is reversed, although in these cases self-similarity is less clear. As the self-similarity is stronger for eastward forcing, and consequently eastward mean flow, we infer that this self-similar behaviour is influenced by the emergent mixing barrier that forms in the present set of numerical experiments. Additionally, by relating the scaling factor of this self-similarity to the global enstrophy of the flow, we have related the dependence of the mean and standard deviation of potential vorticity to the global balanced quantities of the flow in a statistically steady state. Importantly, this reduces the problem of determining the mean and standard deviation to only needing to know a single quantity: the enstrophy (or eddy enstrophy in the case of the standard deviation). Further work will be needed to fully determine the range of parameter space for which this self similarity behaviour exists. It is also clear that the implications of this result have not been fully elucidated with respect to eddy-mean flow interaction in a statistically steady state jet. In future work it is important to relate this self-similar scaling to the imposed parameters of the model such as the wind stress strength.

The behaviour of skewness and kurtosis as a function of time-averaged streamfunction is characteristic of a mixing barrier and a simple statistical model was used to paint a statistical picture of a mixing barrier. Any potential stochastic parameterization must respect the multimodal, and mixed, statistics of a mixing barrier along with its consequent skewness and kurtosis behaviour. This highlights the importance of understanding the influence of highly non-Gaussian noise on the mean flow. It has

been shown that the statistics is no longer multimodal when viewed from a frame of reference moving with the barrier itself. This suggests a separation between the statistics due to the diffusive eddies in the homogenized regions of potential vorticity and the statistics of the barrier meandering, that is, the statistics of the large scale Rossby waves. This also highlights the potential importance of Lagrangian thinking when attempting to parameterize eddy statistics.

We have argued that the results presented here have important implications for stochastic parameterizations as well as being an useful tool for the study of ocean jets and their influence on mixing. Despite the increased complexity of comprehensive ocean models, the statistical picture we have painted relies on the existence of a mixing barrier rather than on the barotropic nature of the dynamics. For this reason, we anticipate that our results should have much broader applicability to baroclinic flows (cf. Hughes et al., 2010) it will be important to relate the neutral density dependence of the statistics to the depth dependence of the statistics; we expect strong mixing barriers forming near the surface but not at depth, consistent with observations of the Gulf Stream (Bower et al., 1985), observation-constrained data in the Southern Ocean (Abernathey et al., 2010), and idealized numerical calculations (Esler, 2008a,b). The extension of these ideas to a realistic ocean model is the subject of Chapter 5.

In conclusion, we have presented a simple yet useful description of the statistics of a barotropic jet with a mixing barrier over a wide range of forcing and dissipation strengths. The jet meandering can be described by Bernoulli statistics while the diffusive turbulence either side is described by a superposed small-scale noise. This way in which the mixing barrier separates the statistical populations of potential vorticity either side can be considered not only as a statistical model of mixing barriers but as a expression of what a mixing barrier *is* in a statistical sense. We hope that the statistical picture of mixing barriers presented here can provide a foundation for future work on the stochastic representation of eddies and their dynamics in ocean

models.

Chapter 5

Statistical perspectives on the Southern Ocean

5.1 Introduction

In this chapter I will consider how the statistics considered in Chapter 4 can teach us about potential vorticity fronts, jets and mixing barriers in a realistic representation of the Southern Ocean. As discussed in Chapter 1, the Southern Ocean plays a key role within the global ocean circulation and eddies play a singular role, as it is only through eddies that large scale Southern Ocean dynamical balances are maintained. It is therefore pertinent to understand the interactions between these turbulent motions and the larger-scale Antarctic Circumpolar Current.

As a requisite to understanding the potential vorticity structure of the Antarctic Circumpolar Current it is important to locate potential vorticity fronts and to determine their nature. In order to pursue this goal it important to carefully define the terminology used, as discussed in more detail in Section 2.7, but repeated here for clarity:

- A *potential vorticity front* is a region of high potential vorticity gradient sepa-

rating regions with differing potential vorticity.

- A *mixing barrier* is a potential vorticity front which prohibits mixing across the front.

- A *jet* is a region of high mean velocity, often well aligned with the zonal direction.

As we found in Chapter 4, where jets were found both with and without mixing barriers; we find that in the complicated potential vorticity structure of the Southern Ocean, it is important to make the distinction between the three terms above.

In this study we consider the utility of potential vorticity statistics in understanding the location and nature of potential vorticity fronts diagnosed from the Southern Ocean State Estimate (SOSE). Further, we consider the relationship of these statistics to measures of eddy diffusivity in order to understand the influence of mean flows on the eddies. The specific aims of this chapter are as follows:

- To use skewness/kurtosis, as in Chapter 4, to locate potential vorticity fronts within SOSE by using simple statistics viewed in various frames of reference.

- To interpret the relationship of these potential vorticity fronts to mixing barriers and jets.

- To test the relationship between the kurtosis of potential vorticity and eddy diffusivity.

A summary of the chapter is as follows. In Section 5.1.1 a short literature review is given on identifying jets and potential vorticity fronts in the Southern Ocean. In Section 5.1.2, a short literature review is given on the suppression of mixing by mean flows in the Southern Ocean. In Section 5.2, SOSE data is introduced and described. In Section 5.3, simple statistics of the Southern Ocean flow are computed at the surface, at depth and on neutral density surface, the implications of these statistics for

fronts, jets and mixing barriers are discussed. In Section 5.4, a calculation of isopycnal (neutral density) Osborn-Cox (eddy) diffusivity is presented and its relationship to the simple statistics is discussed. In Section 5.5, distributions of potential vorticity on neutral density surfaces are considered. In Section 5.6, some concluding remarks are given.

5.1.1 Potential vorticity front identification

There have been various other methodologies proposed to identify Southern Ocean jets, three of which are introduced below:

- *Gradient threshold methods.* Here jets are searched for by looking for locations where the sea surface height gradient exceeds some threshold (Graham et al., 2012; Chapman, 2014).

- *Contour methods.* Jets are located by fitting contours which pick out continuous contours of high sea surface height gradient (Volkov and Zlotnicki, 2012).

- *Probability distribution function methods.* Jets are located by looking for spatially bimodal distributions of sea surface height (Thompson et al., 2010), similar to Hughes et al. (2010) who consider temporal statistics. We expect a sea surface height to be bimodal in the region of jets as high sea surface height gradient, or a step, is related to high velocity through geostrophic balance.

Both gradient threshold methods and contour methods rely on taking the gradient of ocean fields which reduces the effective resolution as well as amplifies noise. Contour methods however, have the advantage of unifying a dynamical and hydrographical perspective of ocean jets. Probability distribution function methods are simpler but do produce a spatially or temporally coarse-grained picture of ocean jets. For a full review of the literature and the advantages and disadvantages of each method the reader is referred to the extensive discussion in Chapman (2014).

While the method of Hughes et al. (2010) is not the only method of jet detection from sea surface height data and is perhaps not the most effective (one major disadvantage, for example, is that it relies on temporal statistics and therefore does not allow for tracking jet locations in time), it has the advantage of relying on simple potential vorticity temporal statistics, allowing for an examination of the flow at depth and in the complex geometry of isopycnal surfaces.

5.1.2 Mixing suppression by mean flows

Mixing by mesoscale eddies in the Southern Ocean has received much attention (Marshall et al., 2006; Shuckburgh et al., 2009b; Abernathey et al., 2010; Sallée et al., 2011; Naviera Garabato et al., 2011; Abernathey and Marshall, 2013) and, in particular, the question of the how mean flows suppress this mixing (Ferrari and Nikurashin, 2010; Klocker et al., 2012).

Bower et al. (1985), found that the Gulf Stream suppresses mixing in the surface flow while permitting mixing across the jets at depth. The paradigm of Dritschel and McIntyre (2008) is that the inhomogeneous nature of geostrophic turbulence leads to self rectifying jets and the formation of a potential vorticity front[1] which prohibits mixing across it. Ferrari and Nikurashin (2010) and Klocker et al. (2012) argue that as the background flow decays with depth it reaches a level where zonal velocity becomes equal to the Rossby phase speed leading to the breakdown of the mixing barrier.

5.2 The Southern Ocean State Estimate

Due to the sparsity of in-situ observations, we calculate the statistics from the Southern Ocean State Estimate (SOSE) (Mazloff et al., 2010), a numerical model at 1/6°

[1]Often referred to as a 'potential vorticity front' or the 'potential vorticity staircase'.

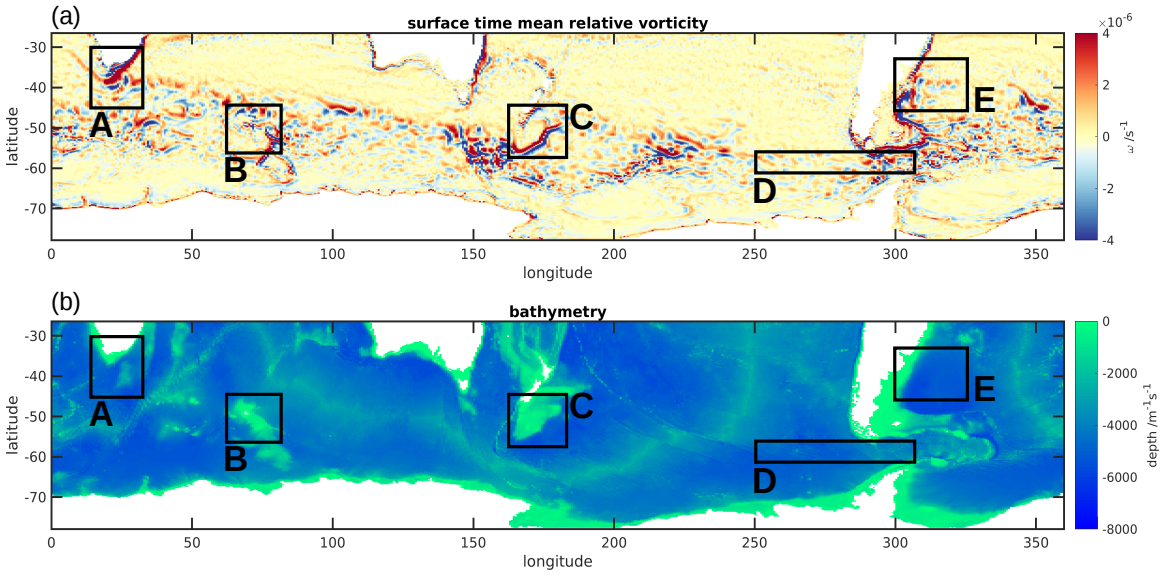


Figure 5.1: a) Time-mean surface relative vorticity and b) bathymetry (depth) of the Southern Ocean. The locations of geographical regions which are relevant to this study: A, the region of the Agulhas current retroflection (as it encounters the Antarctic Circumpolar Current) with Agulhas rigs propagating north-west from the region towards the South Atlantic; B, the region of the Kerguelen Plateau; C, the region of the Campbell Plateau; C, the approximate region of the DIMES experiment (Gille et al., 2007), passing through the Drake Passage; E, the region of the Malvinas western boundary current and its extension.

resolution, constrained to fit, as closely as possible, a wide range of observational data which include: altimetry observations of sea surface height; ARGO float profiles; as well as observations from elephant seals. It is important to note that SOSE data is *not* observational data; SOSE data is derived from a numerical ocean model and subject to the caveats normally associated with ocean models, such as issues with numerical convergence. There are many sources of uncertainties contained within SOSE, including structural errors in the model, resolution errors in the model as well as observational measurement errors. However, SOSE is found to be an improved representation of the observational data compared to the often used *World Ocean Atlas 2001* (Mazloff et al., 2010).

Figure 5.1a shows the time-mean surface relative vorticity characterizing the typical flow found from SOSE. In order to contextualize this flow geographically the

bathymetry is given in Figure 5.1b.

Hughes et al. (2010) showed that the skewness and kurtosis of relative vorticity can be used to identify strong jets from altimetry observations; we will compare the results of Hughes et al. (2010) to the skewness/kurtosis diagnosed from SOSE surface relative vorticity.

Using data from the SOSE we can also look for this skewness/kurtosis signatures at depth. This will be done in both a Eulerian[2] sense and in a frame of reference following time-varying neutral density surfaces (Jackett and McDougall, 1997), a means of producing a surface on which motions are approximately adiabatic. Isopycnal potential vorticity dynamics is analogous to ideal quasi-geostrophic dynamics, used in Chapters 4 and 6, as they are both statements of the material balance of potential vorticity between its sources, sinks. We will take statistics on 5-day mean neutral density surfaces of the Ertel potential vorticity

$$Q = \frac{(\boldsymbol{\omega} + \mathbf{f}) \cdot \nabla \gamma_n}{\rho_0}, \tag{5.1}$$

where $\boldsymbol{\omega}$ is the three-dimensional relative vorticity, $\mathbf{f}$ is the planetary vorticity, γ_n is the neutral density, and ρ_0 is the reference density under the Boussinesq approximation. By calculating the potential vorticity from Southern Ocean State Estimate data and interpolating onto 5-day mean neutral density surfaces we are able to compute the skewness and kurtosis as functions of longitude, latitude in both depth and neutral density coordinates. The time period of the integration is from 1$^{\text{st}}$ of January 2005 until 31$^{\text{st}}$ of December 2010.

[2]At a fixed point in longitude, latitude and depth.

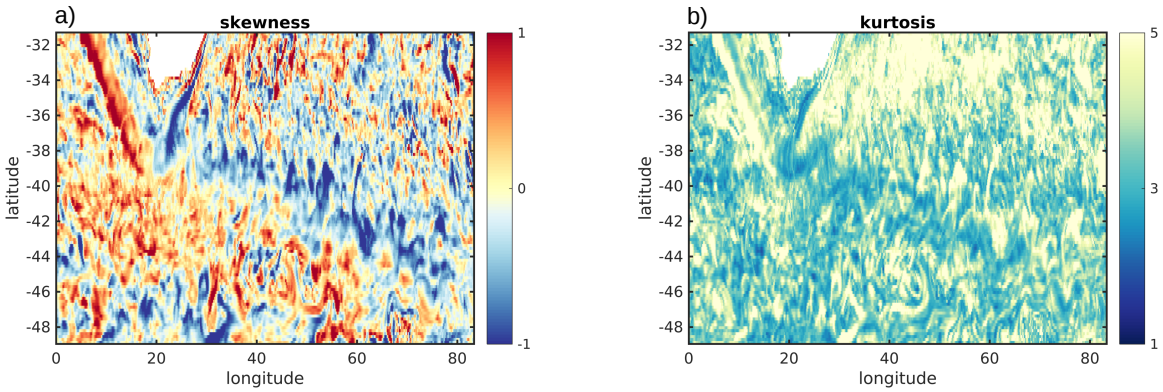


Figure 5.2: a) Skewness and b) kurtosis of surface relative vorticity (de-trended and with seasonal cycle removed) in the region of the Agulhas current retroflection for the whole time period.

5.3 Simple potential vorticity statistics: skewness and kurtosis

5.3.1 Surface statistics

We begin by considering the statistics of the surface flow in order to draw a direct comparison with the work of Hughes et al. (2010). Figure 5.2a and b give the skewness and kurtosis of the surface relative vorticity respectively; the data have been de-trended and the seasonal cycle has been removed by subtracting the mean for each month. Shown, is the region of the Agulhas retroflection where the fronts can be seen in the simple statistics.

Following Hughes et al. (2010) and Chapter 4 we look for positive-to-negative skewness feature concomitant with a high-low-high kurtosis feature as an indicator of a meandering potential vorticity front. As per Hughes et al. (2010), we argue that relative vorticity, as well as potential vorticity, will exhibit a front in the core of a jet. The skewness/kurtosis features are identified by inspection according to the criteria: 1) a skewness feature is a strong negative (approaching $S = -1$) next to a strong positive (approaching $S = 1$) while a kurtosis feature should be a platykurtic region

($1 < K < 3$) flanked by leptokurtic regions ($K > 3$); 2) the skewness and kurtosis features are concomitant; 3) a feature is extended in space, predominantly in a zonal direction; 4) a feature is consistent with our expectation of where strong currents should be. This methodology is somewhat subjective and this should be considered a caveat for the work presented in this chapter. In future work, it would be advantageous make this procedure automatic but this is unnecessary for the current study where we are interested in understanding the underlying statistics from a physical perspective.

The resulting skewness/kurtosis maps display a noisy structure and it not possible to locate all the expected Antarctic Circumpolar Current fronts from the simple diagnostic. A weak signal can, however, be identified to the south of South Africa in the vicinity of the Agulhas Current retroflection (Figure 5.1 and Figure 5.2). Three features with positive-negative skewness and high-low-high kurtosis are worth noting: 1) a strong and narrow signal extending southwards from where the Agulhas Current detaches from topography; 2) a weaker signal in the retroflection region where the Antarctic Circumpolar Current and the Agulhas current merge; and 3) a strong and clear signal associated with the formation of Agulhas rings as they drift into the South Atlantic. This third skewness/kurtosis feature, associated with coherent eddies rather than meandering fronts, is not discussed in Hughes et al. (2010) nor in Chapter 4, and indicates the important influence that various coherent structures in geophysical flows have on higher order statistics of the flow.

By identifying the jets in the Agulhas retroflection region we have been able to reproduce the result of Hughes et al. (2010) who find the strongest skewness/kurtosis signal here. The skewness/kurtosis diagnostics of Hughes et al. (2010) give a stronger (and cleaner) signal for the Agulhas Current retroflection, as well as identifying further jets within the Antarctic Circumpolar current using satellite altimetry data (Le Traon et al., 1998). There are a few possible explanations for the discrepancy. Although

the altimetry data only has a horizontal resolution roughly equivalent to 1°, we must remember that the SOSE data is from a numerical simulation (constrained by data) in an eddy *permitting* regime; by not fully resolving the mesoscale ocean turbulence we expect that the fronts within the SOSE simulation may not be as sharp and that the fronts may exhibit reduced meandering compared with observational data. In addition, the SOSE time-series is six years long, substantially shorter than the available observational data from altimetry. It is also not clear what the effect is of including modifying surface boundary conditions, via optimization, as is done for SOSE. The modified surface forcing changes the potential vorticity at the surface likely leading incorrect flow structures.

5.3.2 Eulerian statistics

SOSE data allows us to consider potential vorticity statistics directly (i.e. not simply relative vorticity statistics) as well as consider the depth structure of these statistics. Diagnosing the potential vorticity defined by equation (5.1), we can compute the Eulerian skewness and kurtosis at depth; that is, statistics derived from the point-wise time-series in an Eulerian coordinate system. Figure 5.3 shows various statistics of pertinent quantities associated with the Antarctic Circumpolar Current. Skewness is not shown here for brevity as it is found that in all places where there is a high-low-high kurtosis feature, a concomitant positive-negative skewness feature is present.

Figure 5.3a shows a horizontal cross-section of kurtosis at a depth of 328m. Here we see two kurtosis (high-low-high) features, one south of Australia and another west of the Drake Passage, indicating the presence of a meandering potential vorticity front. Vertical cross-sections of these kurtosis are given in Figure 5.3b and c (longitudes, 143.4° and 288.2°) displaying the depth structure of this feature. By comparing with the time-mean zonal velocities, Figure 5.3f and g, at the same longitudes we see that the kurtosis feature is concomitant with a region of high time-mean zonal velocity:

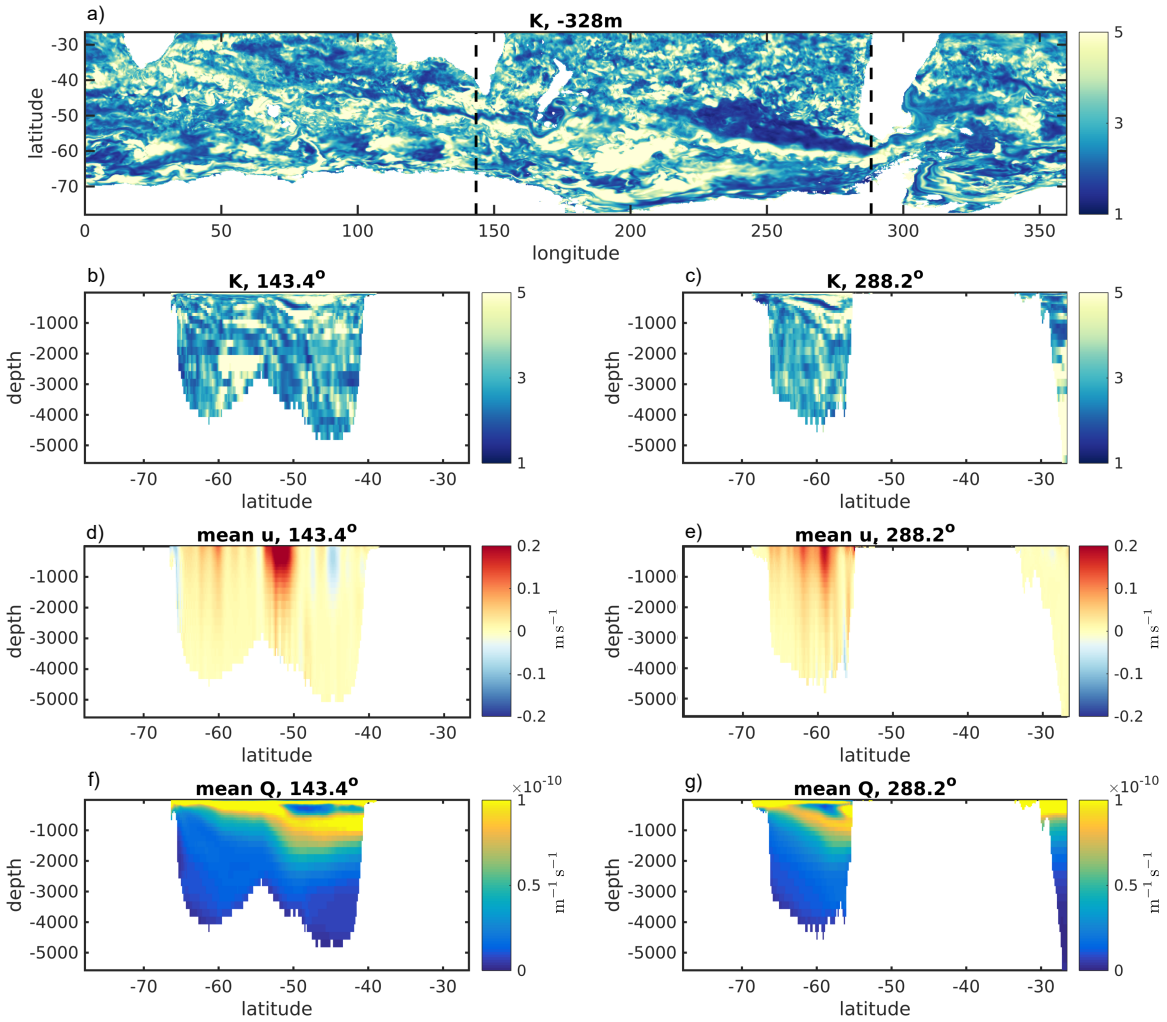


Figure 5.3: a) A horizontal cross-section of potential vorticity kurtosis at depth $z = -328$m. b) A vertical cross-section of potential vorticity kurtosis at longitude of 143.4°. c) A vertical cross-section of potential vorticity kurtosis at longitude of 288.2°. d) A vertical cross-section of time-mean zonal velocity at longitude of 143.4°. e) A vertical cross-section of time-mean zonal velocity at longitude of 288.4°. f) A vertical cross-section of time-mean zonal potential vorticity at longitude of 143.4°. g) A vertical cross-section of time-mean zonal potential vorticity at longitude of 288.2°.

a jet. The kurtosis feature is only present near surface whereas the time-mean zonal velocity extends to depth and has an equivalent barotropic nature; that is, while not formally barotropic the flows are characterized by structures with a deep vertical extent. To establish that the kurtosis feature is picking out a meandering potential vorticity front we plot the time-mean potential vorticity, shown in Figure 5.3f and g. We see that there is indeed a potential vorticity front between a high potential

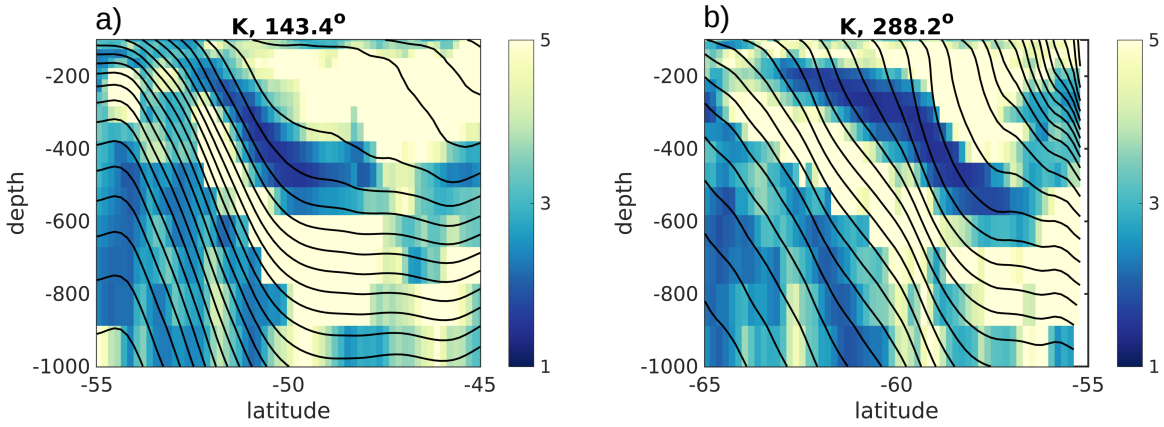


Figure 5.4: Detailed plot of high-low-high kurtosis feature in a vertical cross-section at longitude a) 143.4°and b) 288.2°. Contours of time-mean neutral density are shown.

vorticity tongue coming down from the surface and a region of low potential vorticity to the north of the jet. The low region of potential vorticity is associated with the formation of mode water, a weakly stratified water mass formed by a deep winter mixed layer which is capped by the shallower summer mixed layer (Hasson et al., 2012).

Similarly to in Chapter 4, we find a high-low-high kurtosis signature associated with a jet (region of high time-mean zonal velocity) and a potential vorticity front. We now consider whether this kurtosis feature indicates a barrier to adiabatic mesoscale eddy mixing. In order to do this we consider the kurtosis feature in more detail. Figure 5.4 gives a closer image of the kurtosis for both latitudinal cross-sections and overlaid with contours of the time-mean neutral density. We interpret such a feature, following Chapter 4 and Hughes et al. (2010), as a property of a moving potential vorticity front as the kurtosis is indicative of a bimodality in the full time series of potential vorticity at each point in space. It is apparent that the high-low-high kurtosis feature is aligned approximately along the neutral density contours, especially for Figure 5.4a (144.4°), and implies that the potential vorticity front and its motion is perpendicular to the time-mean neutral density surfaces. That is, the motion of the potential vorticity front is dominated by the motion of the isopycnal surface itself.

This alignment with the time-mean neutral density surfaces explains why these skewness/kurtosis features cannot be seen in the statistics of depth integrated flow; the normal to the neutral density surface is closely aligned to the vertical (due to the ocean's aspect ratio), the potential vorticity front is lost when an integral in depth is taken.

Importantly, the fact that the potential vorticity front is in a direction perpendicular to the time-mean isopycnal excludes the possibility that these potential vorticity fronts are also mixing barriers. Mesoscale eddy mixing is, to a good approximation, adiabatic (e.g. Iselin, 1939; Montgomery, 1940); therefore, the potential vorticity front needs to be seen in isopycnal coordinates in order to imply a prohibition of mesoscale eddy mixing along the surface.

5.3.3 Neutral density statistics

When potential vorticity statistics are viewed in an Eulerian frame of reference we have shown that it is possible to identify meandering potential vorticity fronts at depth associated with jets. However, it not possible to associate these features with mixing barriers. To examine whether the simple statistical diagnostics considered here can be used to identify mixing barriers as well as potential vorticity fronts we will consider statistics on the instantaneous neutral density surface in this section.

Maps of skewness and kurtosis are given for two neutral density surfaces in Figure 5.5. Figure 5.5a and b have several features of this type (highlighted in boxes) and corresponds to regions of high potential vorticity gradient within the Antarctic Circumpolar Current as well as the Malvinas current (east of South America, Figure 5.1). Although weaker in other places the skewness/kurtosis feature seems to extend through the whole domain. Looking at Figure 5.5c and d, this is more apparent; now a connected feature can be seen circumnavigating the Southern Ocean. The feature in Figure 5.5c and d, has shifted south compared to Figure 5.5a and b; the southward

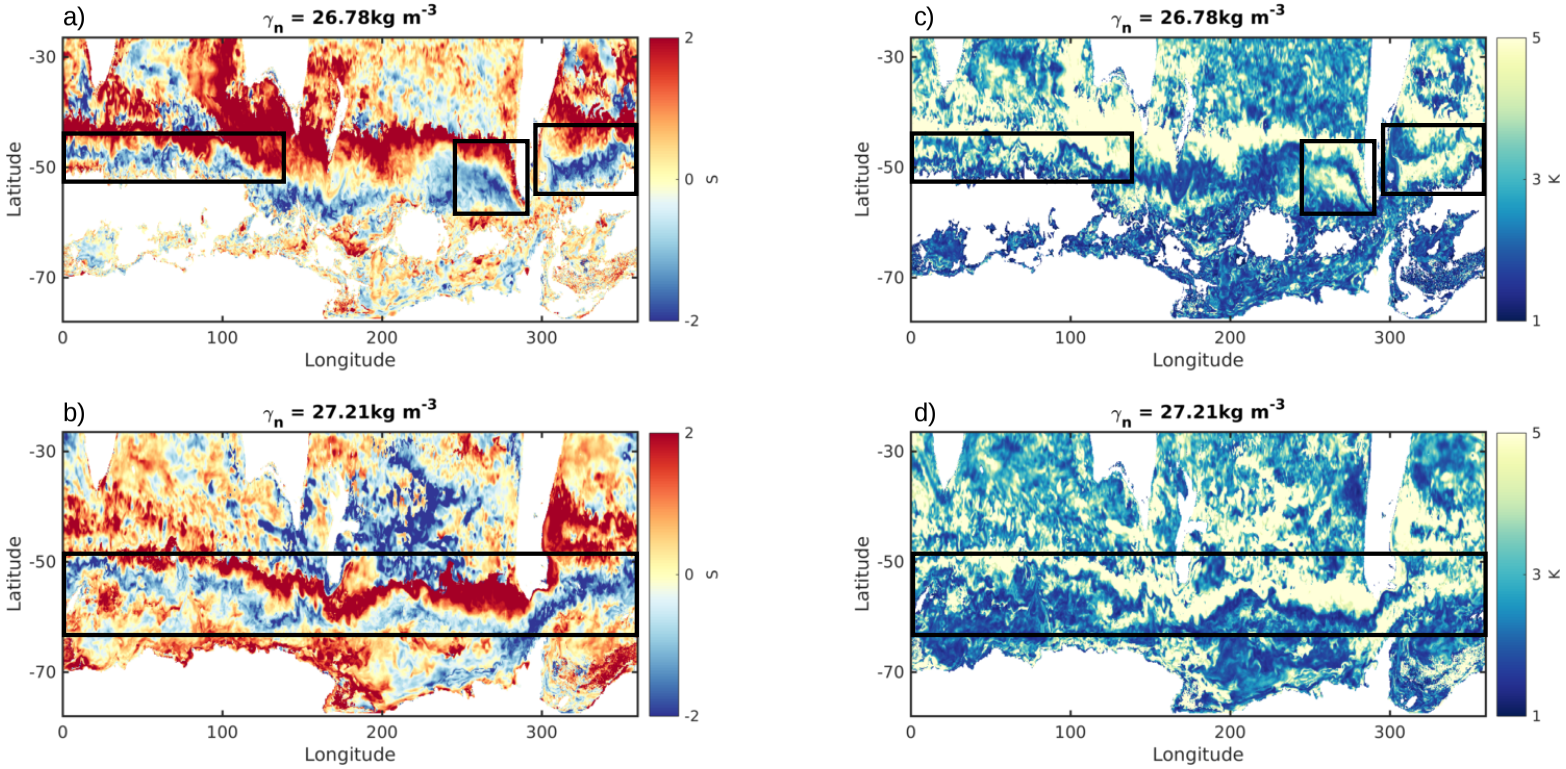


Figure 5.5: Maps of skewness and kurtosis for two neutral density surfaces from Southern Ocean State Estimate data. a) Skewness at $\gamma_n = 26.78$. b) Kurtosis at $\gamma_n = 26.78$. c) Skewness at $\gamma_n = 27.21$. d) Kurtosis at $\gamma_n = 27.21$. Boxes indicate regions of interest. We can see positive-negative features in the skewness concomitant with high-low-high features in the kurtosis. These can be interpreted as meandering potential vorticity fronts corresponding to a mixing barrier. *White* - regions of no data due to outcropping or land.

migration of the skewness/kurtosis feature can be seen very clearly when additional density surfaces are examined.

Cross-sections of the kurtosis and skewness in density coordinates are shown in Figure 5.6. From this we can see the skewness/kurtosis feature is global, being extant for most latitudes. Figure 5.7a and b compare the summer/winter isopycnal potential vorticity for neutral density, $\gamma_n = 27.2\text{kg}\,\text{m}^{-3}$; when the evolution of potential vorticity is inspected on the neutral density surface, it is observed that as the outcrop moves from north (winter) to south (summer), low potential vorticity is deposited while a potential vorticity front remains at the location of winter outcropping throughout the course of the year, until the isopycnal outcrops the following winter. For each neutral density surface, the sharp drop-off in kurtosis north of the longitudinally-extended high-low-high feature (see Figure 5.5b) corresponds to the most northerly point of outcropping; this is shown in Figure 5.7c where the kurtosis is plotted with regions, which outcrop at any point in the SOSE time-series, masked out. When location which outcrop are masked out we lose the distinctive high-low-high kurtosis feature which was seen in Figure 5.5b. We can, therefore, conclude that the strong circumpolar high-low-high kurtosis feature seen in Figure 5.5 is associated with seasonal and inter-annual variability of outcropping and the formation of high potential vorticity at the base of the mixed layer. Once again potential vorticity kurtosis is able to identify a potential vorticity front; however, this front cannot be associated with a jet embedded within the Antarctic Circumpolar Current but with with seasonal and inter-annual mixed layer dynamics.

5.4 Osborn-Cox diffusivity

Interestingly if we examine the region roughly considered by the DIMES experiment (Gille et al., 2007), see Figure 5.1, we find a remarkable qualitative similarity between

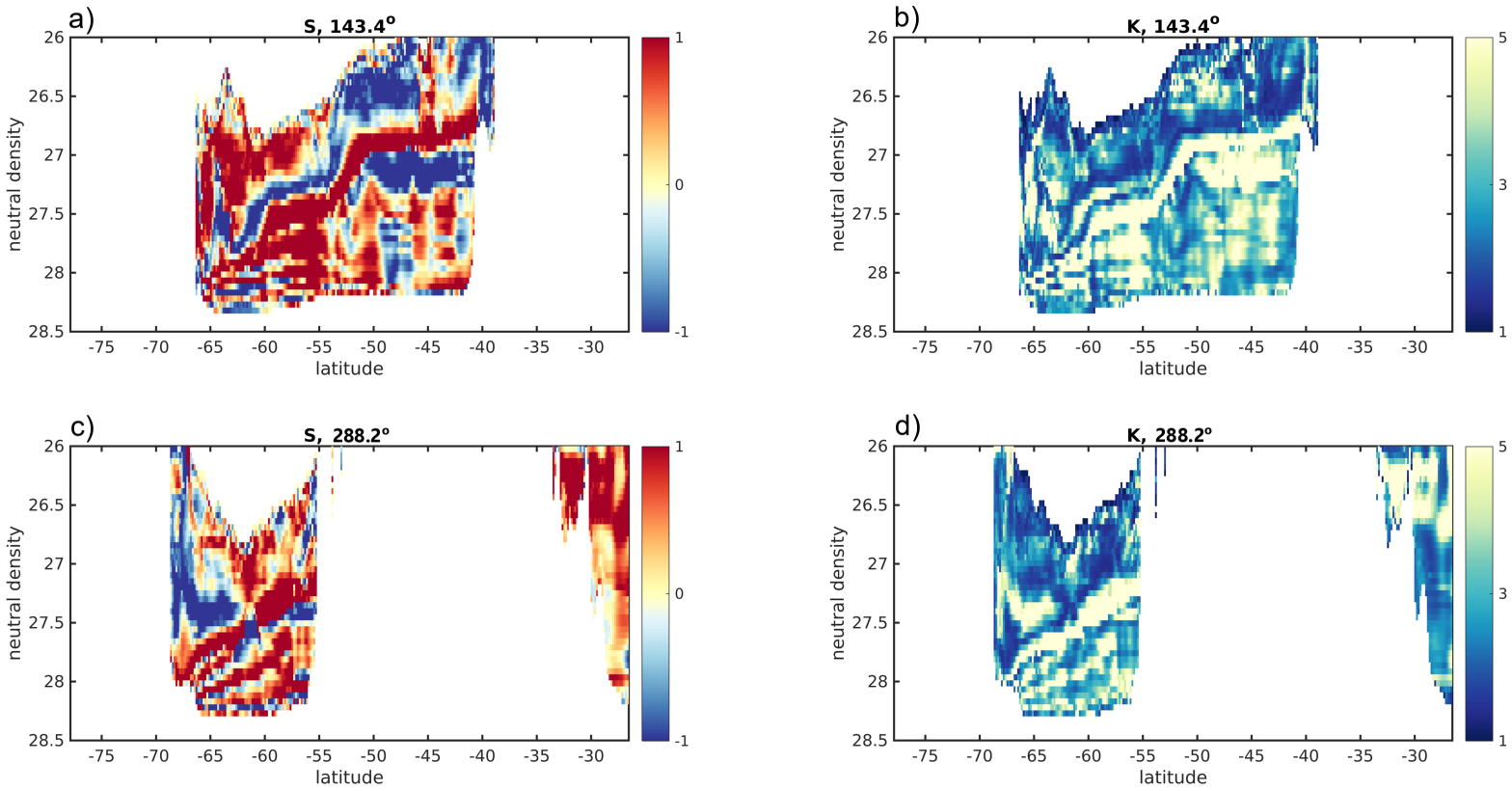


Figure 5.6: a) A vertical cross-section in neutral density coordinates of skewness at longitude of 143.4°. b) A vertical cross-section in neutral density coordinates of kurtosis at longitude of 143.4°. c) A vertical cross-section in neutral density coordinates of skewness at longitude of 288.2°. d) A vertical cross-section in neutral density coordinates of kurtosis at longitude of 288.2°.

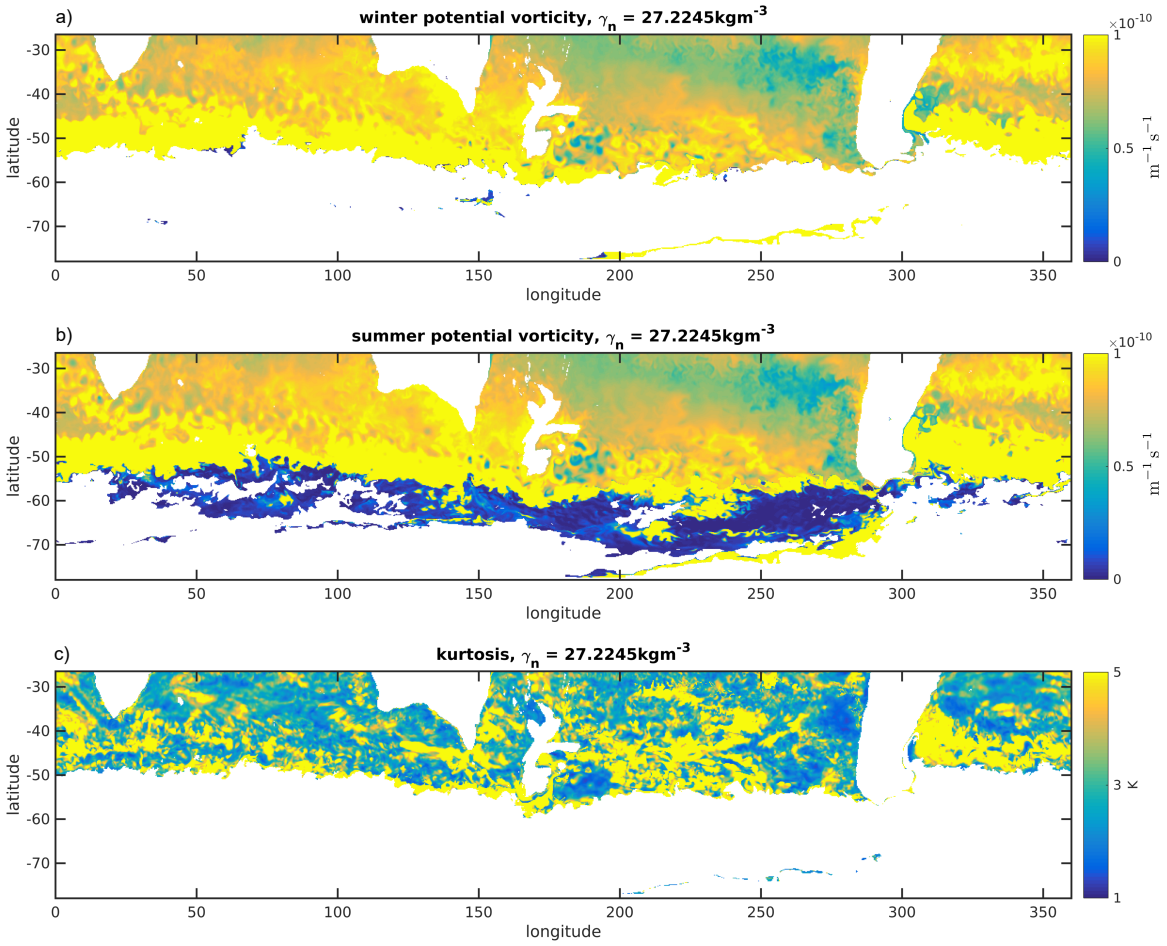


Figure 5.7: a) Typical snapshot of potential vorticity on neutral density surface, $\gamma_n = 27.2\text{kg m}^{-3}$, during winter. b) Same is plotted during the subsequent summer (6 months later). c) Map of kurtosis where all points which outcrop at any time during the time-series have been masked out. Now, we do not see the longitudinally-extended high-low-high feature.

the kurtosis signal seen in Section 5.3.3, caused by seasonal and inter-annual variation of the outcropping, and the isopycnal effective diffusivity, a Lagrangian measure of eddy enhancement of down-gradient diffusion, as reported by Abernathey et al. (2010) for the $\gamma_n \approx 27.2\text{kg m}^{-3}$ neutral density surface. This suggests that one needs to carefully consider the influence of the seasonal variation in neutral density surfaces before assigning dynamical causes for the regions of enhanced and reduced eddy diffusivity. In this section we consider an Eulerian measure of eddy diffusivity to quantify the nature of mixing on the neutral density surface.

Many studies (e.g. Marshall et al., 2006; Shuckburgh et al., 2009a,b; Abernathey et al., 2010; Ferrari and Nikurashin, 2010; Sallée et al., 2011; Naviera Garabato et al., 2011; Klocker et al., 2012; Abernathey and Marshall, 2013) have used various methods to quantify eddy mixing in the Southern Ocean. In Hughes et al. (2010) a correlation was found between the kurtosis of potential vorticity and effective diffusivity (Nakamura, 1996) in a two layer quasi-geostrophic model. It is important therefore to consider eddy diffusivities in order to test the picture formed from our analysis of skewness and kurtosis in the preceding sections.

In this study we will compare the diagnosed isopycnal kurtosis to the Osborn-Cox diffusivity (Osborn and Cox, 1972; Abernathey and Marshall, 2013) of potential vorticity. We choose Osborn-Cox diffusivity, instead of effective diffusivity, for its: 1) local nature, making it a more appropriate comparison to Eulerian kurtosis; and 2) the comparative ease of computation allowing us to be confident with the results with a complex geometry, complex topography as well as with a complex structure in the potential vorticity distribution (see Section 5.5). The Osborn-Cox diffusivity is related to effective diffusivity through an integral along a tracer contour (Nakamura, 1996); the Osborn-Cox diffusivity is interpreted as an Eulerian measure of eddy diffusivity.

The form of the Osborn-Cox diffusivity can be found by considering the tracer variance budget in the Eulerian-mean framework and separating out the part of the eddy-flux which is down-gradient and associated with irreversible mixing and is found to be

$$\kappa_{oc} = \kappa \frac{\overline{|\widetilde{\nabla} Q'|^2}}{|\widetilde{\nabla} \overline{Q}|^2}, \tag{5.2}$$

where κ is the small-scale potential vorticity diffusivity used in the model; tildes denote the two-dimensional gradient taken along the neutral density surface; and the over-line denotes the isopycnal time-mean while the primes denote deviations from this mean. For a full derivation see, for example, Nakamura (2001).

In order to compute the Osborn-Cox diffusivity we relate the isopycnal two-

dimensional gradient, $\widetilde{\nabla}$, to derivatives in the horizontal and vertical

$$\widetilde{\nabla}Q(x,y,\gamma_n) = \nabla Q(x,y,z)|_{z=z_\gamma} + \left.\frac{\partial Q(x,y,z)}{\partial z}\right|_{z=z_\gamma} \nabla z_\gamma(x,y), \tag{5.3}$$

where z_γ is the depth of the neutral density surface, and ∇ denotes the horizontal two-dimensional gradient. The correction due to the curvature of the isopycnal is small but we include this in the computation for completeness. For a primitive equation model, such as SOSE, it is difficult to determine the value of κ as potential vorticity is diagnosed and not a prognostic variable. We therefore define the *eddy enhancement factor*, ϵ, as

$$\epsilon = \frac{\kappa_{oc}}{\kappa} = \frac{\overline{|\widetilde{\nabla}Q'|^2}}{\overline{|\widetilde{\nabla}\overline{Q}|^2}} \tag{5.4}$$

which is a non-dimensional number which gives the factor by which eddies enhance the small-scale diffusivity. This factor can be interpreted as a ratio between the local length-scales of eddy-perturbed to time-mean contours which leads to a measure of the efficiency of mixing (Abernathey and Marshall, 2013).

Figure 5.8 compares the eddy enhancement factor, ϵ, with the kurtosis for neutral density surface, $\gamma_n = 27.2\mathrm{kg\,m^{-3}}$. We find low values ($\epsilon = 1$, no eddy enhancement of mixing) in regions associated with jets and coherent structure: paths of Agulhas rings; Agulhas Current retroflection; and Malvinas Current extension. The enhancement of mixing is, on the whole, higher in the region south of the most northerly outcropping, see Figure 5.7c.

It can be seen that the regions of low ϵ do not occur where the strongest high-low-high kurtosis signal can be found (associated with the seasonal and inter-annual variability of outcropping) but rather occurs further north in the core of the Antarctic Circumpolar Current; this is consistent for all values of neutral density. Weaker high-low-high kurtosis features (longitude 300°to 350°,latitude -50°to -40°) can seen which are concomitant with low mixing enhancement, for Figure 5.8 this is particularly clear

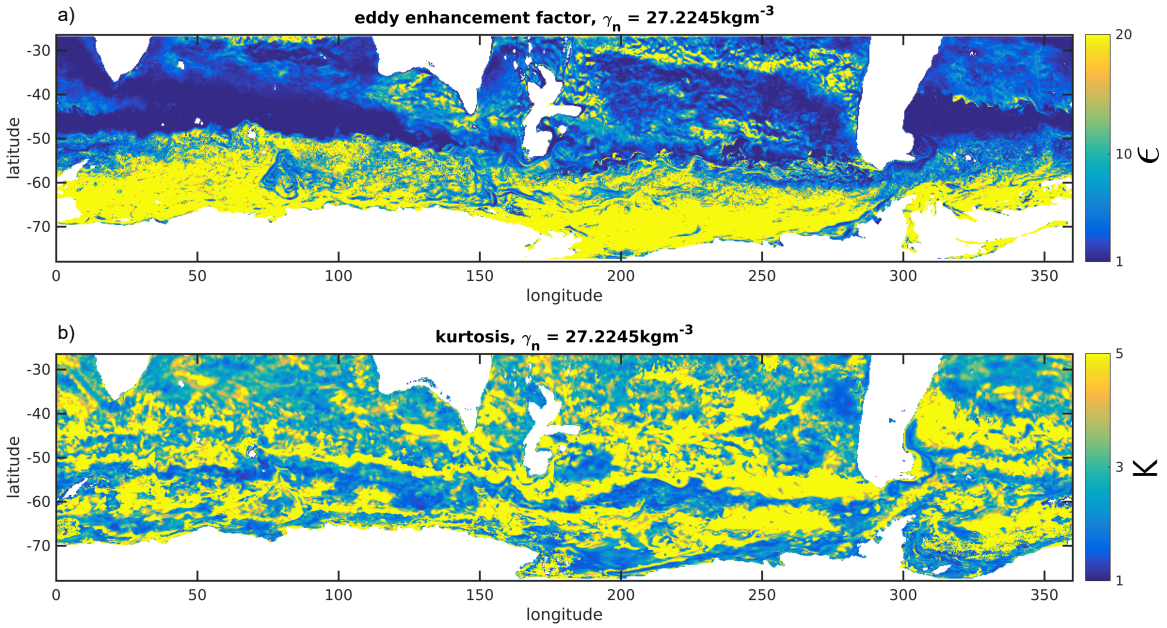


Figure 5.8: a) Map of the eddy enhancement factor ϵ on a neutral density surface. b) Map of kurtosis of potential vorticity on a neutral density surface. The neutral density surface for a and b is $\gamma_n = 27.2\text{kg m}^{-3}$.

in the Malvinas Current extension.

5.5 Distributions of potential vorticity

5.5.1 Global distributions of potential vorticity on neutral density surfaces

To further understand the nature of the potential vorticity front presented in Section 5.3.3 we turn to global potential vorticity distributions on the instantaneous neutral density surface. Figure 5.9 shows histograms of potential vorticity on four neutral density surfaces, $\gamma_n = (26.71, 26.96, 27.22, 27.48)\text{kg m}^{-3}$. The histograms use all temporal and spatial data for each surface. We find that when we split the data into depth classes of $z > -200$m and $z < -200$m, consistent with a typical value of the winter mixed-layer depth in SOSE, we see a marked distinction in the shape of the histograms. This result suggests that the global potential vorticity statistical popu-

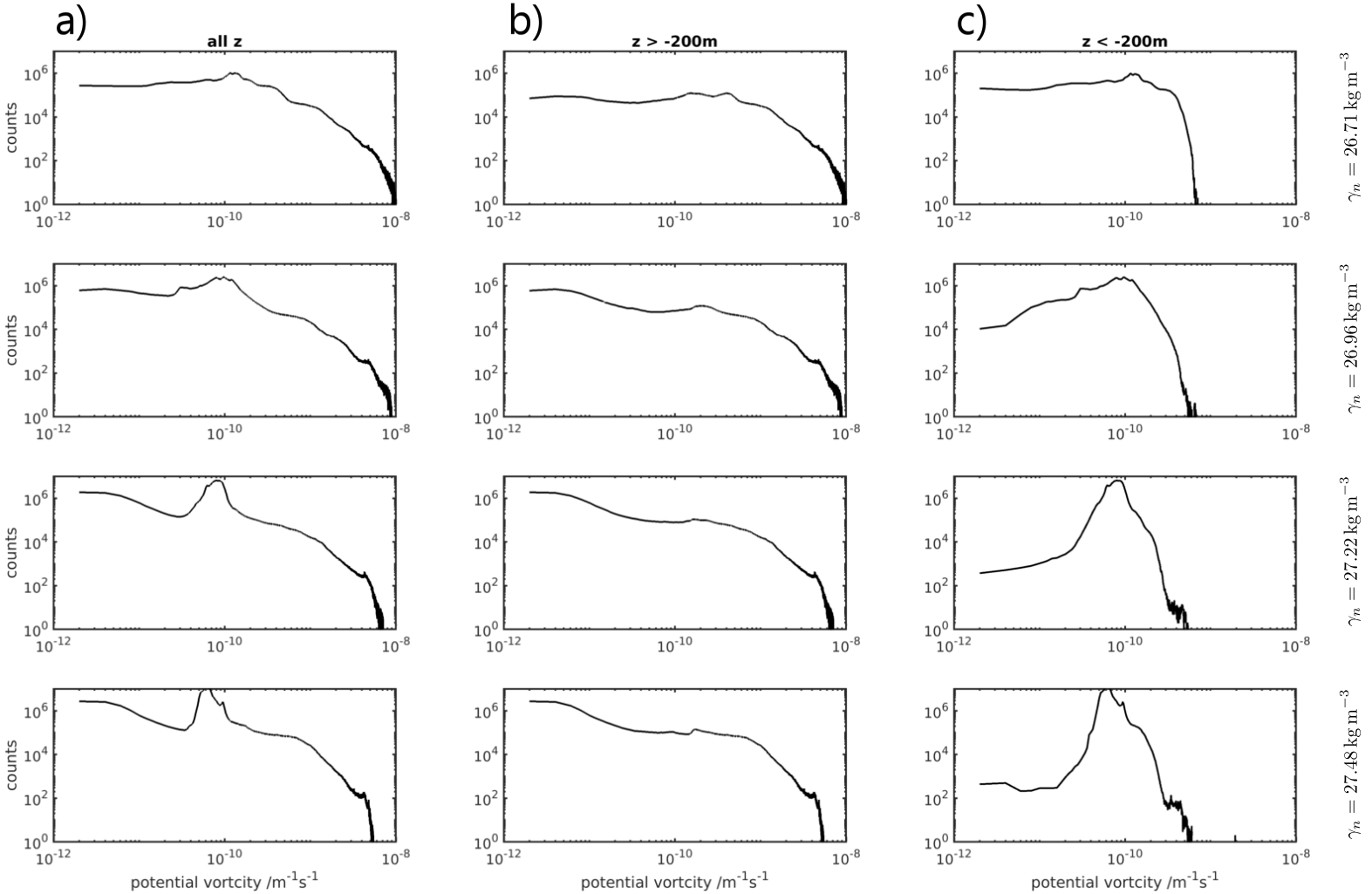


Figure 5.9: Global histograms of potential vorticity for the neutral density surfaces, $\gamma_n = (26.71, 26.96, 27.22, 27.48)\mathrm{kg\,m^{-3}}$. Histograms for the a) full statistics as well as b) the $z > -200$m and c) the $z < -200$m sub-populations.

lation can be separated into sub-populations implying that the distributions shown in Figure 5.9 can be best described as mixed distributions such as those described in Section 4.7.2.

The distributions shown in Figure 5.9 are considerably different from those presented in Chapter 4, with the two statistical sub-populations which have drastically different statistics. The statistics of potential vorticity for $z > -200$m have a broad structure ranging from $10^{-12}\,\mathrm{m^{-1}s^{-1}}$ to of order $10^{-8}\,\mathrm{m^{-1}s^{-1}}$, there is a peak at $Q = 10^{-12}\,\mathrm{m^{-1}s^{-1}}$ with a plateau region extending over a number of orders of magnitude where the counts are approximately uniform (plots are on a log-log scale).

This should be contrasted with the distribution of the $z < -200$m sub-population of potential vorticity. Here we see a much narrower distribution with the potential vorticity peaking slightly below $10^{-10}\,\mathrm{m^{-1}s^{-1}}$, which is consistent across a wide range of neutral density surface. The asymmetric tail towards lower potential vorticity, after experimentation with various choices of sub-sampling in depth, is found to be associated with an imperfect separation into independent sub-populations of the potential vorticity statistics.

When we consider the variation between the neutral density surfaces we see that the separation of the two sub-populations of potential vorticity becomes less distinct on lighter surfaces. An asymmetric low potential vorticity tail persists in the lower sub-population ($z < -200$m), suggesting that the potential vorticity does not effectively mix all the way down the neutral density surface. The low potential vorticity tail in the $z < -200$m distribution becomes more pronounced with decreasing neutral density while the cut off at high potential vorticity remains a robust feature for all values of neutral density. For the $z > -200$m sub-populations the distribution is flatter for lighter surface, consistent with a picture of enhanced surface mixing.

In Chapter 4 we found mixed sub-populations of homogenized potential vorticity statistics leading to mixed probability distributions (Section 4.7.2). Is this same

behaviour seen in the potential vorticity statistics considered here? In order to understand the implications of mixing on a neutral density surface we consider the following thought experiment in the long-time asymptotic limit of strong diffusive mixing.

Consider the diffusive mixing of potential vorticity in two regions. Region A has imposed values of potential vorticity at the boundaries, Q_{LOW} and Q_{HIGH}; region B has no imposed values of potential vorticity at the boundaries but has a no-flux condition at the boundary. The temporal evolution of potential vorticity in each region, A and B, is given by

$$\frac{\partial Q}{\partial t} = \nabla \cdot (\kappa \nabla Q). \tag{5.5}$$

If we assume a spatially constant diffusivity κ, the long time spatial distribution is a solution of Laplace's equation

$$\nabla^2 Q = 0. \tag{5.6}$$

Due to the differing boundary conditions the solution will be very different in the two regions, schematically represented in Figure 5.10. In region A, the *spatial* distribution of potential vorticity will become linear and the corresponding *probability* distribution will become an uniform distribution. In region B, the *spatial* distribution of potential vorticity will become uniform and the corresponding *probability* distribution will become a Dirac-delta function.

We can use the thought experiment described here to qualitatively understand the distribution in the two regions, $z > -200$m and $z < -200$m, on the neutral density surface. The upper region of the neutral density surface, $z > -200$m, should be compared with region A in the thought experiment. At the surface of the ocean we expect the potential vorticity to be low due to very small vertical gradients of the neutral density, $\partial_z \gamma_n \approx 0$, while at the base of the (winter) mixed-layer we expect the potential vorticity to be be very high as the vertical neutral density gradients become very high due to the interface between the unstratified mixed-layer and the

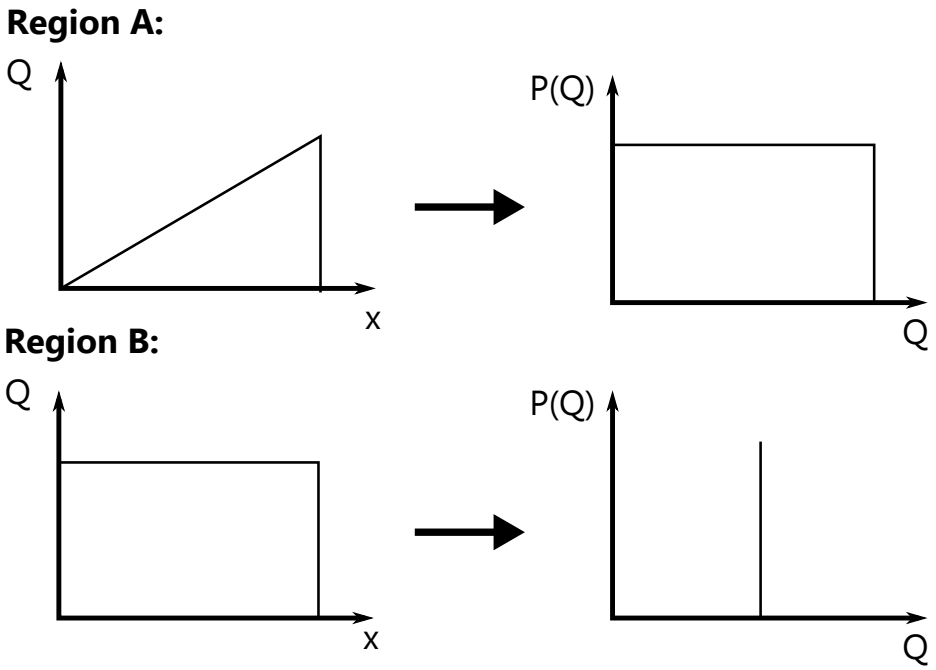


Figure 5.10: Spatial distribution of potential vorticity and the corresponding probability distribution under strong diffusive mixing for region A, with prescribed boundary values, and region B, with no prescribed boundary values.

stratified abyssal ocean. In summer, the outcrop recedes south leaving the region of high potential vorticity and an associated front, see Figures 5.3.3 and 5.7. Near the surface however, the potential vorticity must remain low. This sets up imposed values of potential vorticity at the surface and base of the $z > -200$m region, determined by the large-scale stratification and mixed layer dynamics. We can see the effect of diffusive mixing in the flat plateau in the histograms plotted in Figure 5.9.

The lower region of the neutral density surface, $z < -200$m, should be compared with region B in the thought experiment. Away from the influence of the mixed layer we expect the spatial distribution of the potential vorticity to be homogenized, one value of potential vorticity for all spatial points, as in region B of the thought experiment. We can see this tendency in Figure 5.9: with increasing neutral density the $z < -200$m sub-population becomes more homogenized as the distribution becomes narrower.

Thus, in a very different way to Chapter 4, the distributions in Figure 5.9 represent

mixed distributions, formed from two well-mixed sub-populations of the potential vorticity statistics, where the higher distribution is spread out between the boundary values of potential vorticity and the deeper region is homogenized in the sense of Chapter 4.

5.5.2 Geographic variation in potential vorticity distributions

To gain more insight into the isopycnal potential vorticity statistics at depth, $z < -200$m, we consider two-dimensional histograms of potential vorticity jointly with longitude, $n(Q, x|\gamma_n)$, latitude, $n(Q, y|\gamma_n)$, and depth, $n(Q, z|\gamma_n)$; these are shown in Figure 5.11 for the same four neutral density surfaces as in the previous section.

Looking at the plots of $n(Q, z|\gamma_n)$ and $n(Q, y|\gamma_n)$, Figure 5.11, we see increased homogenization and a narrowing of the distribution with depth. This is consistent as you look at the denser neutral density surface as well as when you descend down to deeper locations on the surface.

The histogram with longitude, $n(Q, x|\gamma_n)$, shows the same effect, with the spread in values of potential vorticity decreasing with an increase in the value of neutral density. It is also possible to understand particular oceanographic features by comparing Figure 5.11 to the potential vorticity on the neutral density surface. The Drake Passage is clearly seen in the histograms as a 'gap' in the statistics where the ocean becomes narrow. It is also possible to associate regions where the distribution are skewed to higher potential vorticity values with the Agulhas Retroflection and the Malvinas Extension; regions where the distributions are skewed to low potential vorticity values can be associated with regions of mode water. The statistics often exhibit transitions, for example a positive to negative skew, at topographical features such as the Campbell Plateau and the Drake Passage (Figure 5.1).

The joint histograms given in Figure 5.11 confirm the conclusions of the similar analysis of Thompson et al. (2010) that topography plays an important role in

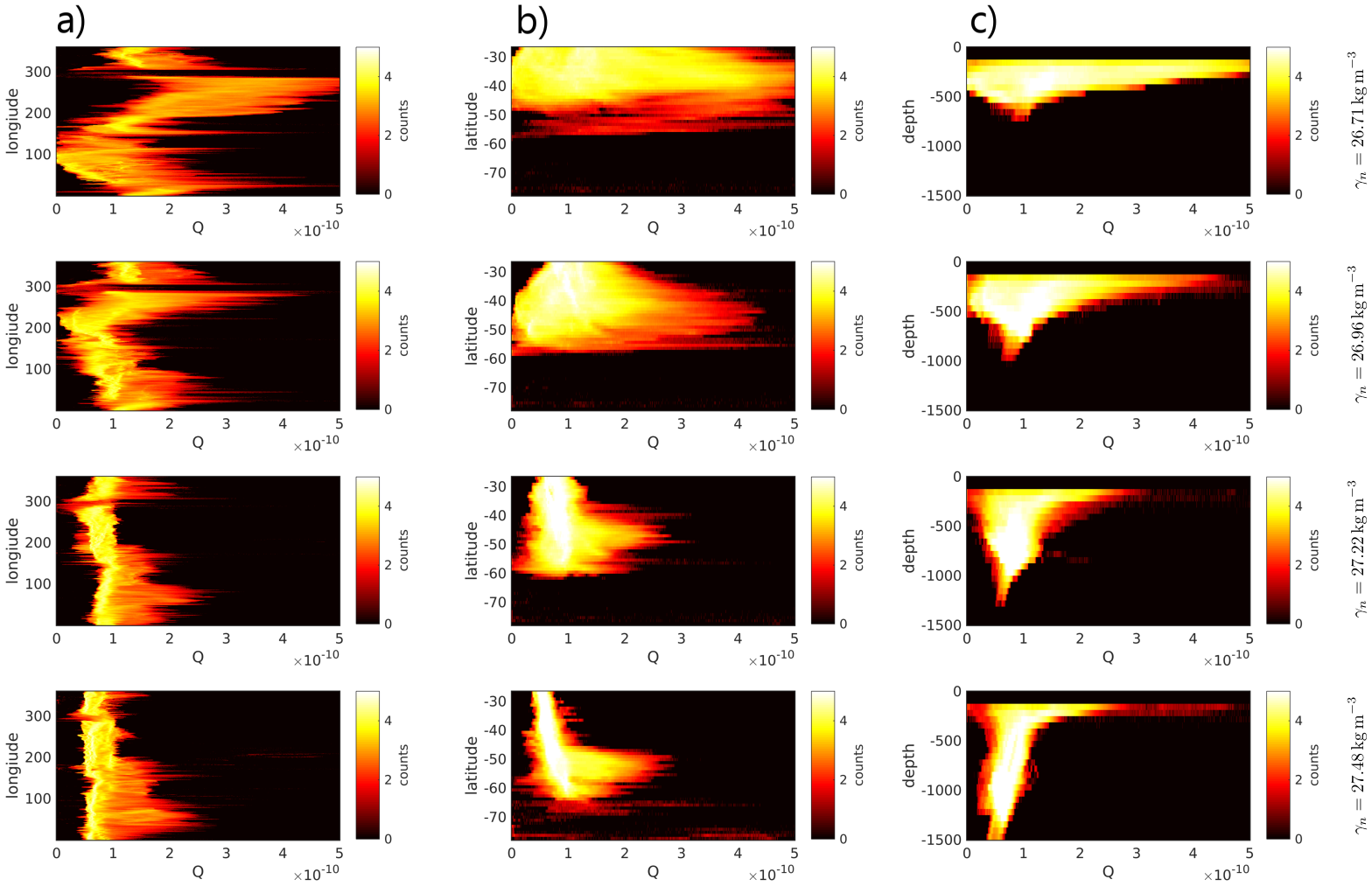


Figure 5.11: Joint histograms of potential vorticity and an ordinate for the neutral density surfaces, $\gamma_n = (26.71, 26.96, 27.22, 27.48)\mathrm{kg\,m^{-3}}$. a) Joint histogram, $n(Q, x|\gamma_n)$. b) Joint histogram, $n(Q, y|\gamma_n)$. c) Joint histogram, $n(Q, z|\gamma_n)$.

determining the structure of potential vorticity distributions.

5.6 Concluding remarks

In the results presented in Section 5.3 it is shown that the skewness/kurtosis of potential vorticity (or relative vorticity in the case of the surface flow) can be used to qualitatively locate potential vorticity fronts in the Southern Ocean. We can determine the nature of these fronts by inspecting the character and structure of the flow. It is found that the identified potential vorticity front can be classified in three ways.

- *Potential vorticity fonts which are jets and mixing barriers.* These features can be seen in both the surface statistics of relative vorticity, Section 5.3.1, and the neutral density statistics, Section 5.3.3. The signals from the skewness/kurtosis are weaker than those found in Hughes et al. (2010) but are consistent topographically fixed features; for example, the Agulhas Current retroflection and the Malvinas Current extension.

- *Potential vorticity fronts which are jets but are not mixing barriers.* When the statistics are calculated at fixed depth it is not possible to infer a mixing barrier as stirring from geostrophic eddies is adiabatic to a good approximation. When we compare the observed skewness/kurtosis feature with time-mean neutral density surfaces, Figure 5.4, it becomes clear that the identified potential vorticity fronts are not barriers to geostrophic mixing.

- *Potential vorticity fronts which are not jets but are potentially mixing barriers.* When we consider the statistics on neutral density surfaces we see another type of potential vorticity front at the base of mixed layer. The characteristic skewness/kurtosis signal is associated with seasonal and inter-annual variability

in the outcropping of the neutral density surface. Although this front is related to the base of the mixed layer and not Antarctic Circumpolar Current jets, it is possible that this front is nevertheless a mixing barrier, prohibiting the mixing of potential vorticity down the neutral density surface.

As well as considering simple statistics of the potential vorticity, the Osborn-Cox diffusivity was calculated on neutral density surfaces. It was found that there is a qualitative similarity between regions of high and low kurtosis and regions of high and low eddy enhancement. This is particularly clear were we find topographically fixed jets (Malvinas extension, Agulhas retroflection) as well as in the track of Agulhas rings propagating into the South Atlantic. Within the main portion of the Antarctic Circumpolar current, however, the region of low eddy enhancement is found immediately north (or below) the potential vorticity front situated at the base of the winter mixed-layer. This suggests that the Antarctic Circumpolar Current provides a cap contributing to the persistence of this potential vorticity front during summer as the outcrop retreats southward.

Further evidence is found for a partial barrier to mixing associated with the bottom of the mixed-layer when we consider the global distribution of potential vorticity on neutral density surfaces. It is found that the full statistics can be partially decomposed into two sub-populations classified by depth, $z > -200$m and $z < -200$m, with the decomposition improving for denser (and deeper) neutral density surfaces. It is argued that each of the two sub-populations can be thought of as well-mixed with the very disparate distributions coming from constraints on the potential vorticity due to stratification structure of the Southern Ocean. This rather crude analysis and identification of sub-populations could be improved by an unsupervised classification method (e.g. Maze et al., 2017) although the highly non-Gaussian nature of the $z > -200$m statistics will make a Gaussian Mixture Model approach unfeasible.

When the joint distributions of potential vorticity with longitude, latitude and

depth ($z < -200$m) on neutral density surfaces are considered, we find that the potential vorticity becomes more homogeneous at depth and that topographic features play an important role in determining the potential vorticity structure of the Antarctic Circumpolar Current.

The results presented in this study highlight the complexity of Southern Ocean potential vorticity structure, the importance of the interaction between the mixed-layer and abyssal ocean in considering mesoscale mixing, as well as the the utility of simple statistics in disentangling complex, multi-faceted oceanographic flow.

Chapter 6

Eddy-mixing entropy as a measure of turbulent disorder in barotropic ocean jets

6.1 Introduction

In this chapter we consider the concept of entropy in the context of dissipative and forced-dissipative ocean flow. The powerful theory of equilibrium statistical mechanics is limited in usefulness when tackling non-ideal fluids; incorporating the effect of forcing and dissipation on entropy is essential in addressing the extent to which entropic ideas can be applied to the problem of mesoscale ocean turbulence. Inspired by the Miller-Robert-Sommeria theory we follow here a heuristic approach, akin to that followed in Jung et al. (2006), in order analyse entropy in the context of forced and dissipative numerical simulations. The aims of this chapter are as follows.

- To determine the impact of forcing and dissipation on the evolution of entropy in a turbulent barotropic jet, analytically and numerically.

- To test the maximum entropy principle[1] and to understand the utility of this principle in the context of an idealized forced-dissipative turbulent jet.

- To use the maximum entropy principle as a means to formulate a relationship between dynamically balanced quantities and the statistics of the flow.

In Section 6.2, we describe the numerical experiments. In the previous chapter, Section 3.3.2, we introduce the eddy-mixing entropy and define the key concepts in this study. In Section 6.3, we derive analytical expressions for the influence of constant forcing and linear drag on the entropy as well as consider the case of a Gaussian stochastic forcing. In Section 6.4.1, we diagnose the entropy for a freely-decaying turbulent jet and use this to test the predictions of Section 6.3; in Section 6.4.2, we repeat the diagnosis for a forced-dissipative turbulent jet and consider entropy as a balanced dynamical quantity. In Section 6.5, we consider the maximum entropy principle and relate it to the ideas developed in the subsequent parts of this study. Evidence is given in support of the maximum entropy principle and we discuss this in relation to forced-dissipative flow. In addition the mathematical problem of obtaining the statistics of the flow from the assumption of maximum entropy is presented and the maximum entropy statistics are calculated and compared with simulations. In Section 6.6, we discuss the mean-field approximation for energy and its relation to the maximum entropy statistics. In Section 6.7, the study is concluded with some closing remarks.

6.2 Numerical experiments

Two numerical experiments are performed using the above model. For the details of the model, see Section 2.6.

[1]Not to be confused with the similarly named maximum entropy production principle.

Simulation ID	Wind stress strength, τ_0	Linear drag coefficient, r	Timestep, dt
Freely-decaying experiments			
D_1	0.000	0.0008	0.0005
D_2	0.000	0.0009	0.0005
D_3	0.000	0.0010	0.0010
D_4	0.000	0.0020	0.0010
D_5	0.000	0.0030	0.0010
D_6	0.000	0.0040	0.0010
D_7	0.000	0.0050	0.0010
Forced-dissipative experiments			
FD_1	0.005	0.0050	0.0010
FD_2	0.010	0.0050	0.0010
FD_3	0.020	0.0050	0.0010
FD_4	0.040	0.0050	0.0010
FD_5	0.080	0.0050	0.0010
FD_6	0.160	0.0050	0.0010
FD_7	0.320	0.0050	0.0010

Table 6.1: List of experiments and non-dimensional parameters.

The first experiment is a freely-decaying unstable jet in which the initial jet has a velocity profile

$$u(y) = u_0 \mathrm{sech}^2(y), \tag{6.1}$$

with $u_0 = 10$. The unstable jet evolves freely under the action of hyper-viscosity and varying strengths of linear drag.

The second experiment is a forced dissipative turbulent jet spun up from rest with varied strength of the wind stress. The wind stress profile is kept the same for all simulations as,

$$\tau = \tau_0 \mathrm{sech}^2\left(\frac{y}{\delta}\right), \tag{6.2}$$

but the magnitude of the jet is varied by changing the value of τ_0. δ is the width parameter and is fixed for all simulations. Table 6.1 gives the values used in the different simulations. The parameters which are held constant for both experiments are given in Table 6.2; these parameters were extensively tested for both this study

Parameter	Value
Meridional extent, L_y	$5\pi/2$
Zonal extent, L_x	20π
Number of zonal grid-points, n_x	1024
Number of meridional grid-points, n_y	128
Time-step, dt	1×10^{-3}
Output frequency	1
Total time of output	1×10^{4}
Hyper-viscosity, ν_h	2×10^{-6}
Wind stress width parameter, δ	0.4
Beta parameter, β	0.2

Table 6.2: Fixed parameters of model simulations.

and David et al. (2017) to ensure numerical convergence.

6.3 Analytical model for evolution for entropy

The fundamental quantity that we are interested in is the eddy mixing entropy given by (3.16) in Section 3.3.2. In this section we derive a tendency equation for this entropy. We are not able to derive a full theory as the effects of the non-linear or non-local terms in the vorticity equation, (2.19), Section 2.6, do not seem to be analytically tractable. Nevertheless it is possible to derive analytical expressions for the entropy evolution due to the remaining linear and local terms in the vorticity equation: the wind stress curl and the linear drag.

Ignoring the non-local (hyper-viscous) and non-linear (advection) terms we have the following equation for the evolution of potential vorticity:

$$\frac{\partial q}{\partial t} = -r(q - \beta y) + g(y) \tag{6.3}$$

where $g(y) = -\partial_y \tau(y)$ is the constant (in time) forcing.

Equation (6.3), ignoring the non-local and non-linear terms, leads to the following

partial differential equation for the probability distribution function, ρ,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \tilde{q}}[r(\tilde{q} - \beta y)\rho] - g(y)\frac{\partial \rho}{\partial \tilde{q}}. \tag{6.4}$$

The two terms on the right hand side of equation (6.4) were also derived in Kazantsev et al. (1998) but here we consider the effect of these terms on the entropy. We can write the entropy tendency as

$$\dot{S} = -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\dot{\rho}\ln\rho, \tag{6.5}$$

where the dot represents differentiation with respect to time. By substituting equation (6.4) into (6.5) we can derive (see Appendix D) the influence of these terms on the entropy, yielding

$$\dot{S} = P - Ar, \tag{6.6}$$

where we have now included the effect of advection and hyper-viscosity as the residual, P; and A is the area of the domain. As the hyper-viscosity is small in our numerical calculations, we will take the liberty of referring to P as the *advective production of entropy*.

Notably, (6.6) does not have an explicit dependence on the zonally symmetric forcing as a constant wind stress only shifts the distribution in each zonal band and the entropy is invariant to these shifts. The linear drag leads to a remarkably simple term which is minus the drag coefficient, that is a perpetual and constant sink of entropy. There is an important caveat to the numerical results presented in this study; in a realistic ocean simulation high frequency noise will introduce a source of entropy from its fast time-scale variability analogous to the introduction of a stochastic noise term in the potential vorticity equation; we include the influence of white stochastic noise on the evolution of the probability distribution and entropy in Appendix D.

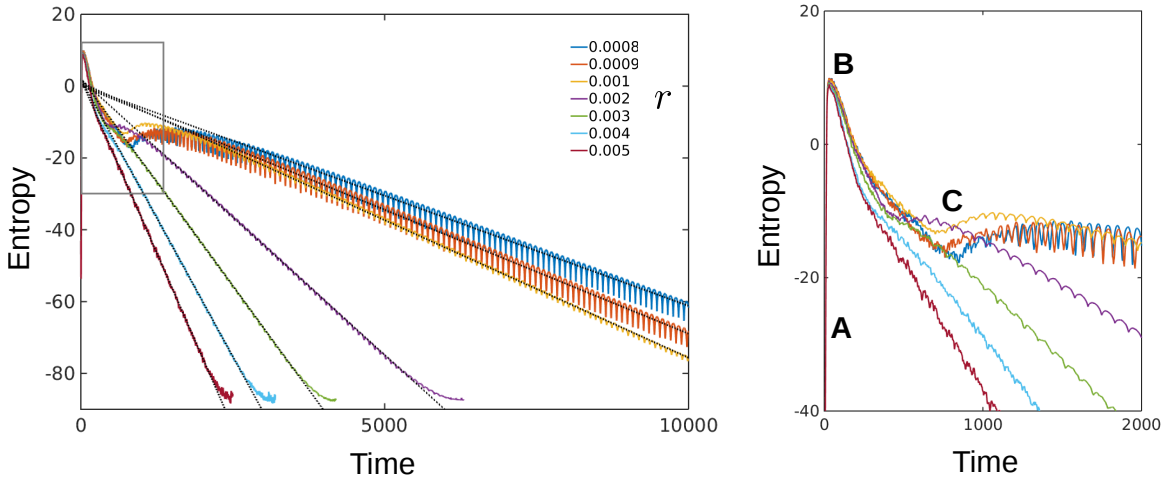


Figure 6.1: Entropy as a function of time for different values of linear drag coefficient, simulations $D_1, ..., D_7$. A - We find that exponential growth of barotropic instabilities lead to a very fast growth in the entropy. B - The entropy reaches a maximum value which is insensitive to the value of the linear drag coefficient. This is followed by a decrease in entropy at a rate greater than the contribution of the linear drag. This implies that this decrease is due to the eddies themselves. C - At long time the entropy decay is linear and is explained by the linear drag. For the lowest values of linear drag coefficient a persistent Rossby wave emerges causing a secondary increase in the entropy before the long-time behaviour is seen.

6.4 Entropy in the numerical simulations

6.4.1 Eddy-mixing entropy in freely-decaying turbulence

As seen in (6.6) the entropy tendency has no explicit dependence on the forcing. However, the forcing will contribute to determining the behaviour of the advective production of entropy, P. To illuminate the effect of forcing we first consider the evolution of entropy in the absence of forcing: freely decaying unstable flow.

We begin by examining the evolution of entropy for short times as the instabilities grow then decay shown in Figure 6.1. In region A of Figure 6.1, the entropy increases very quickly concomitant with the exponential growth of eddy energy through shear instability in the jet. There is little spread in rate of the entropy growth in simulations $D_1, ..., D_7$ with changed drag parameter. This growth is arrested for all experiments at a maximum value in region B for the same time and where the maximum is also not

greatly changed with the differing linear drag coefficient. The entropy then decreases, in the period B to C, towards its asymptotic behaviour. The rate of decrease is *greater* than can be explained by the sink of entropy due to linear drag, Ar. This means that in the period between B to C the advective production of entropy must become negative and acts as a sink of entropy.

There is an interesting difference between simulations $\mathrm{D}_1, ..., \mathrm{D}_3$ and the other simulations. These low drag simulations see a second increase in entropy (Figure 6.1, near Time = 1000) toward the long-time behaviour as well as an oscillation about the long time behaviour. These effects can be illuminated by examining the flow, at low linear drag coefficient a persistent Rossby wave forms causing an increase in the disorder as compared to laminar flow as well as the observed oscillations.

In freely-decaying turbulence the eddies will ultimately die away through the action of linear drag and hyper-viscosity causing the advective production of entropy, P, to tend to zero at long times. In this case equation (6.6) tends to the asymptotic solution for the entropy evolution

$$S(t) \approx -Art + K \quad \text{for long times,} \tag{6.7}$$

where K is a constant of integration. We can test this hypothesis in a simulation of a freely decaying unstable jet as for long times we would expect the activity of the eddies to asymptotically decay to zero. We compare this prediction with the first numerical experiment described in Section 6.2 and shown in Figure 6.1. We see a striking agreement between the long time behaviour, to the right of C, predicted by (6.7) and the slopes diagnosed from the simulations. Figure 6.2 shows the agreement between the predicted and the measured long time slope of the linear entropy decrease which is found to be near exact.

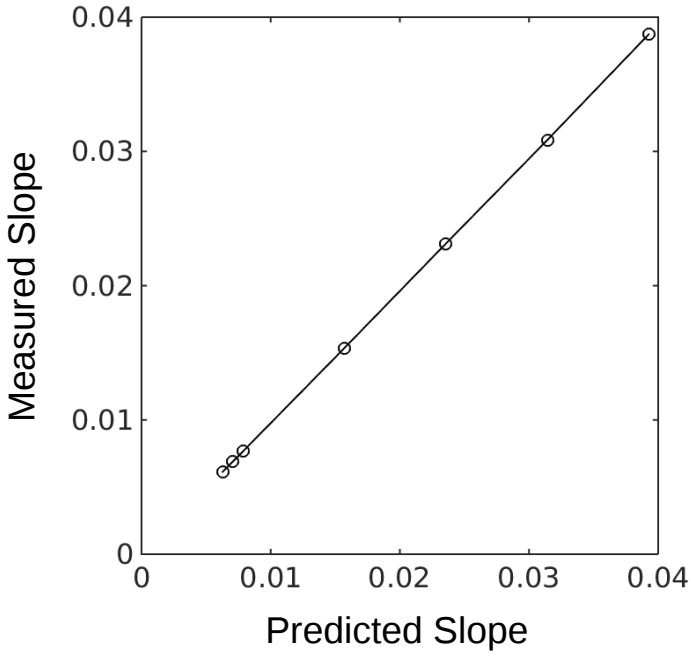


Figure 6.2: Measured slope of long time linear decay of the entropy, diagnosed from simulations $D_1, ..., D_7$, plotted against the predicted slope. The agreement is near perfect.

6.4.2 Eddy-mixing entropy in forced-dissipative turbulence

We now turn our attention to the entropy evolution in the forced-dissipative experiments $FD_1, ..., FD_7$. We start by considering the entropy for short times, comparing it to snapshots of the potential vorticity. As an example we consider experiment FD_3. Figure 6.3 shows the evolution of entropy as the system evolves to a statistically steady state. Initially, in region A, the flow is near laminar with only the small initial perturbation. We see that this corresponds to a low value of entropy. Once instabilities begin to grow the corresponding growth of entropy is fast and grows to a maximum value much like the evolution in the freely-decaying simulations. At the maximum of entropy, region B, the turbulence has covered the whole domain with small scale eddies. As these eddies mix the potential vorticity we see a slump in the entropy. When we examine the flow at the bottom of the slump, region C, we see that a large scale Rossby wave has emerged propagating on a sharp potential vortic-

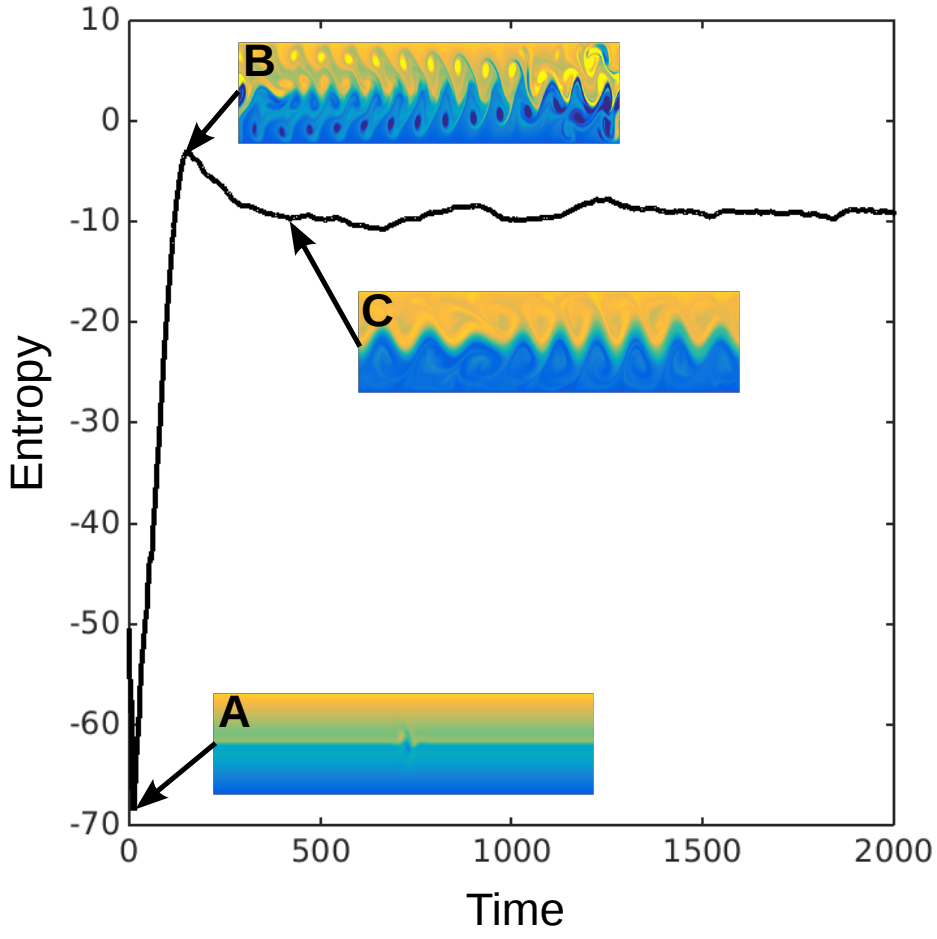


Figure 6.3: Evolution of entropy during spin-up of forced-dissipative flow. A - Initially laminar flow with small perturbation. As the flow becomes unstable the eddy-mixing entropy increases rapidly. B - Entropy reaches a global maximum at the point where turbulence has spread across the whole domain. Disorder is at small scales. C - Coincident with the emergence of a large scale Rossby wave the entropy decreases. Subsequently the entropy fluctuates around a balanced time-mean value.

ity gradient corresponding to a mixing barrier. This decrease of entropy, or disorder, in the system allows us to describe the way in which energy has condensed at large scales in an entropic sense. The concomitance of this decrease in entropy with the emergence of large scales leads to a novel interpretation of well known inverse cascade phenomena: the emergence of coherent large scales can be described as the decrease of entropy in this system.

At longer times the entropy fluctuates around a balanced steady state value. This behaviour of entropy is the same for all the forced-dissipative simulations, $\mathrm{FD}_1, ..., \mathrm{FD}_7$, shown in Figure 6.4a, much like statistically steady state balance of energy, shown in Figure 6.4b. Both the time-mean entropy and energy increase with the wind stress strength in steady state as well as exhibit fluctuations about this mean (although the energy fluctuations are suppressed in Figure 6.4b due to the logarithmic scale). The balanced steady state behaviour of the entropy is explained, according to the reasoning of Section 6.3, by the competition between the advective production of entropy and the constant sink due to linear drag, that is

$$\overline{P} - Ar = 0, \tag{6.8}$$

where the over-line denotes the time-mean in statistically steady state. It is important to note that, because $-r$ is a merely a negative number, both the increase and decrease in the entropy fluctuations arise from the advective production, P. That is, *eddies can act as both a source and a sink of entropy.* It is important to note that the action of P as a source and a sink must be associated with the presence of dissipation which allows there to be fluctuations in otherwise conserved global quantities such as energy and entropy which exhibit inverse and direct transfers between scales.

Although the time derivative of entropy has no explicit dependence on forcing, the forcing does supply energy to the turbulent motions by sustaining the eddy production

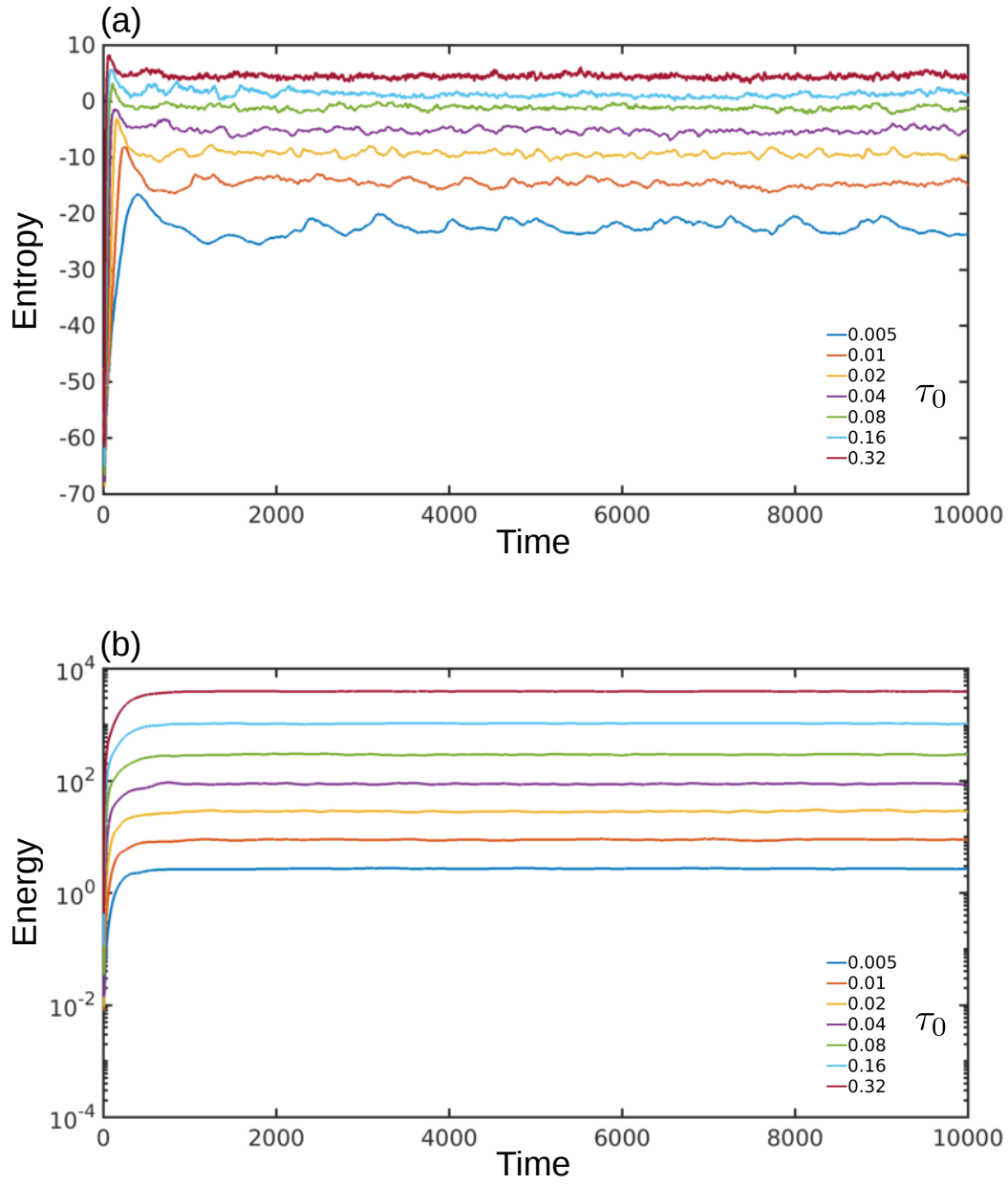


Figure 6.4: (a) Evolution of entropy in simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see that entropy behaves as a balanced quantity in a statistically steady state: in the time mean the sources of entropy are equal to the sinks of entropy. (b) Evolution of energy in simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$ shown on a logarithmic scale. Like entropy the energy is balanced in a statistically steady state. The energy takes a longer time than the entropy to reach steady state balance.

of entropy, unlike in the case of the freely-decaying experiment. The forcing implicitly sets the maximum and steady state value of entropy. We will further consider how the entropy is related to other well-known dynamical quantities in the following section.

6.5 Relation to the maximum entropy principle

6.5.1 Time evolution of entropy and the maximum entropy principle

In what has been discussed so far we have considered the derivative of entropy with respect to time. Now we consider its relation to the maximum entropy principle which is at the core of the equilibrium statistical mechanics theory of ideal geophysical flow (Jung et al., 2006). In order to achieve clarity in our discussion we define the following terminology. For the purposes of this study it is useful to restrictively define *equilibrium* as a maximum entropy state of the ideal flow (i.e. MRS theory). For forced-dissipative systems, we refer to the long-time behaviour as a *(non-equilibrium) statistically steady state*.

The maximum entropy principle states that the entropy should be stationary with respect to variations in the probability distribution, ρ, given appropriate dynamical constraints. We can relate the time derivative of the entropy, S, to the functional derivative using the relation

$$\frac{dS}{dt} = \int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\frac{\delta S}{\delta\rho}. \tag{6.9}$$

Assuming that the system is in a maximum entropy state constrained by the value of

energy and N polynomial Casimirs we have that the variational problem

$$\begin{aligned}\frac{\delta S}{\delta \rho} &+ \alpha(t)\frac{\delta}{\delta \rho}\left(-\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho - E(t)\right) \\ &- \sum_{n=1}^{N}\gamma_n(t)\frac{\delta}{\delta \rho}\left(\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\tilde{q}^n\rho - C_n(t)\right) = 0, \end{aligned} \tag{6.10}$$

is satisfied, where α and γ_ns are Lagrange multipliers and where $E(t)$ and $C_n(t)$ are the time varying values of energy and Casimirs. Substituting this condition into (6.9), we relate the time derivative of entropy to that of energy and the Casimirs:

$$\frac{dS}{dt} = \alpha^*(t)\frac{dE(t)}{dt} + \sum_{n=1}^{N}\gamma_n(t)\frac{dC_n(t)}{dt}, \tag{6.11}$$

where $\alpha^* \equiv -\alpha$, and the detailed derivation of equation (6.11) is given in Appendix E. This equation says that if the entropy is maximal under some constraints then the evolution of the entropy can be entirely explained by the evolution of the quantities constraining it. We can split the time evolution of the Lagrange multipliers into temporal mean and fluctuations such that $\alpha^*(t) = \overline{\alpha^*} + \alpha^{*'}(t)$, and similarly for the other Lagrange multipliers. Assuming that the deviations in the Lagrange multipliers are small, which we will address in more detail in due course, and integrating (6.11), we can write an approximate relation for the time evolution of entropy in terms of the time evolution of energy and the Casimirs, giving

$$S(t) \approx \overline{\alpha^*}E(t) + \sum_{n=1}^{N}\overline{\gamma_n}C_n(t) + K, \tag{6.12}$$

where K is a constant of integration. This expression relies on two assumptions. Firstly, the entropy is maximized constrained by the value of energy and the Casimirs at any point in time; in other words, the system is in a *quasi-equilibrium state* defined as a non-equilibrium statistically steady state where the time scales for changes in

the balanced quantities is larger that the time the eddies take to drive the system to maximum entropy. Secondly, the fluctuations in the Lagrange multipliers (sensitivities of the entropy) are small. We now turn to testing the relation, (6.12), in order to test these assumptions.

6.5.2 Reconstruction of entropy evolution

We can test equation (6.12) by regressing the diagnosed time evolution of entropy onto the time evolution of energy and the other conserved quantities, the Casimirs, and comparing the reconstructed entropy time series against the diagnosed entropy time series. It was found the this procedure does not give a good reconstruction when the spin up is included in the time series, however there is good agreement when we only consider the statistically steady state: it is likely that the time dependence of the Lagrange multipliers is large during spin up but not in the statistically steady state.

An example is given in Figure 6.5 for simulation D_6. Figure 6.5 shows the reconstruction of the entropy time series using only the first two Casimirs, circulation and enstrophy, in addition to the energy as well as using Casimirs up to tenth order. The correlation for the first two Casimirs is 0.73, and when ten Casimirs are used the correlation is 0.93. This is a striking agreement and provides evidence that the turbulence acts to maximize entropy given time-varying constraints, according to equation (6.12), at each point in time in statistically steady state. That is, we find evidence that the system is in a quasi-equilibrium state as defined above. It must be noted that the quasi-equilibrium approximation is likely not to hold during spin-up.

The analysis has been repeated for all forced-dissipative simulations $FD_1, ..., FD_7$ and including differing numbers of Casimirs. Figure 6.6a shows the correlation between reconstructed and diagnosed entropy time series as a function of number of Casimirs for $FD_1, ..., FD_7$. We see a marked increase in the correlation with increased

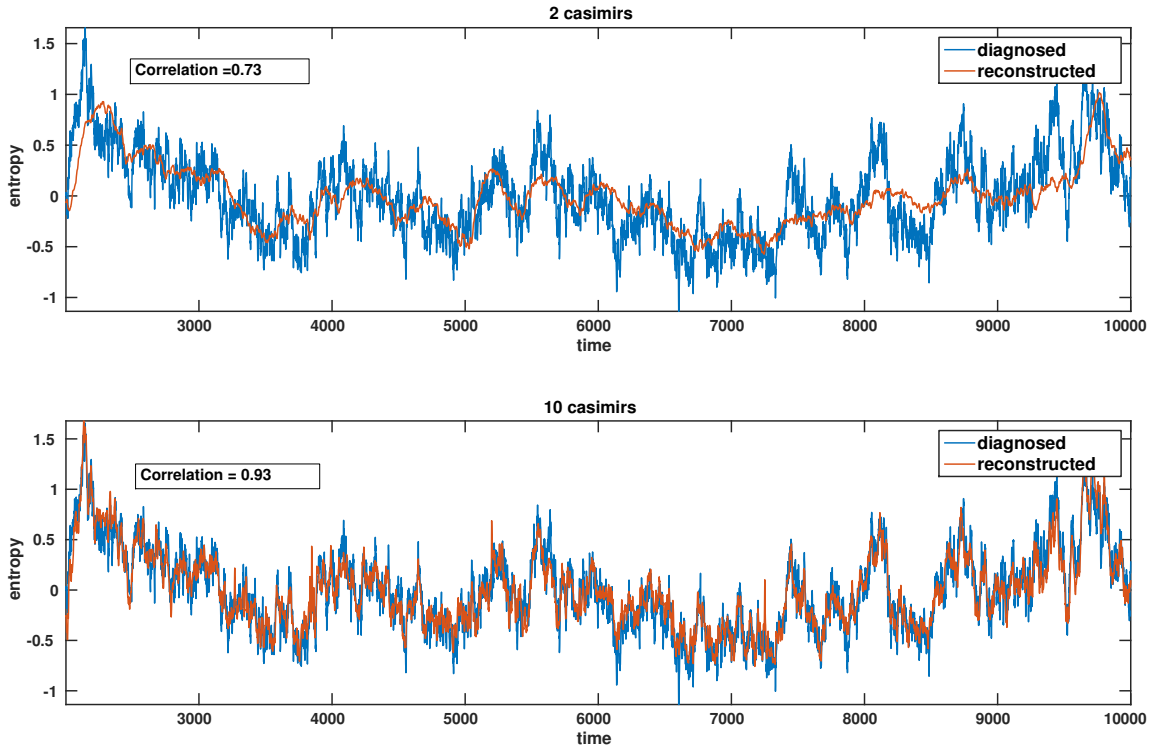


Figure 6.5: Reconstruction of the entropy evolution using relation (6.12). *Top* - using only the first two Casimirs, circulation and enstrophy. *Bottom* - using first ten Casimirs. Both show significant correlation with the two Casimir reconstruction clearly matching the low frequency fluctuations, however but failing to capture the high frequency fluctuations. The ten Casimir reconstruction performs substantially better.

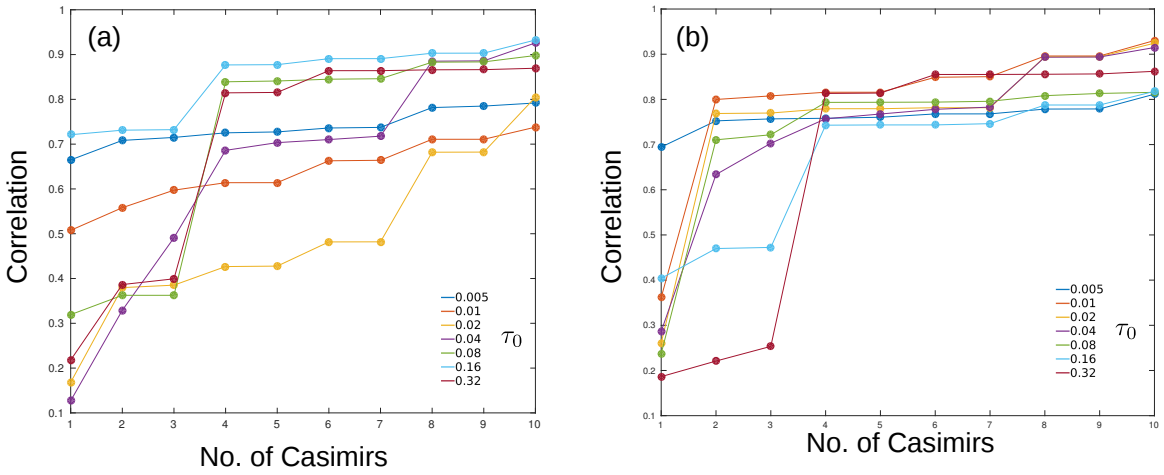


Figure 6.6: (a) Correlation between the diagnosed and reconstructed entropy time series following the relationship (6.12). Entropy is calculated on zonal bands for simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see a clear improvement with number of Casimirs included. This improvement is quantitatively different for each simulation. (b) Correlation between the diagnosed and reconstructed entropy time series following the relationship (6.12). Entropy is calculated on instantaneous streamfunction for simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$. We see a clear improvement with number of Casimirs included. This improvement is quantitatively different for each simulation but now all simulations show high correlation of greater than 0.7 with only four Casimirs. This is to be contrasted with a).

numbers of Casimirs. This shows the importance of higher order Casimirs in this statistical mechanics approach. It is also apparent that odd power Casimirs do not contribute significantly to increasing the correlation. We test the significance of these correlations by comparing with correlations produced by reconstructing the entropy from 'synthetic' time series for the energy and Casimirs. These 'synthetic' time series are produced randomly such that they have the same power spectra as the diagnosed time series. It is found that correlations derived from the model data lie beyond the 99th percentile of an ensemble of 'synthetic' data. In this way we can reject the null-hypothesis that the observed correlations are due to statistical over-fitting. A detailed description of this significance test is given in Appendix F.

We can repeat this analysis, now calculating the entropy along instantaneous streamfunction contours. This means that the macro-cells defined in Section 3.3.2 become contours of streamfunction rather than zonal band as has been used up to now. The entropy will change as a consequence of this transformation but the energy and Casimirs will not - this implies that only the projection in (6.12) of the entropy evolution onto the energy and Casimirs that will change. The correlation of the reconstructed entropy with the diagnosed entropy for this choice of macro-cell is shown in Figure 6.6b.

The main point to note here is that the correlation is higher, and converges quickly, for fewer Casimirs. The fourth order Casimir seems to be of particular importance with all simulations having a correlation of greater than 0.7 if four Casimirs are used. This agrees with the observation of David et al. (2017) that viewing potential vorticity distributions in a more of a Lagrangian sense acts to simplify the statistics requiring a reduced number of moments to describe the local probability distribution of potential vorticity. The particular importance of the fourth order Casimir here is in contrast to the study of Abramov and Majda (2003) who found that, in their numerical set-up, no Casimir above order three was relevant; this indicates that the particular Casimirs

which are most important depend heavily on the particulars of any given system (i.e. domain geometry and nature of forcing).

The fact that it is possible, to a large extent, to reconstruct the statistically steady state time series for entropy using the corresponding dynamically balanced quantities of the flow is tantalizingly suggestive that in a statistically steady state of a forced-dissipative flow the eddies push the system into a quasi-equilibrium: a non-equilibrium statistically steady state which nonetheless maximizes entropy at each point in time with time varying constraints. That is, although not in a true equilibrium (ideal flow with maximal entropy), the rate at which eddies push the entropy to its maximum allowed value is much faster than the time scales over which the conserved quantities are changing and is analogous to the quasi-static approximation in thermodynamics.

As we saw in Sections 6.4.1 and 6.4.2, the eddies can act as a source or sink of entropy. Indeed, the fact that the balanced quantities such as energy (Figure 6.4b) fluctuate at all is due to the turbulence, if there were no non-linearities then we would have steady flow and no fluctuations. We suggest that the eddies play a double role, simultaneously maintaining the quasi-equilibrium and modulating its constraints. Further work, over a wider range of parameters, is required to obtain firmer evidence for the maximum entropy principle at work.

6.5.3 Solving for the Lagrange multipliers

Solving the variational problem, (6.10), gives us a probability distribution in terms of a set of unknown Lagrange multipliers, in this section we describe the method for determining these from knowledge of the energy and Casimirs of the flow. To determine the Lagrange multipliers it is necessary to solve the non-linear simultaneous equations

$$-2E = -\frac{\partial}{\partial \alpha} \int \mathrm{d}^2\mathbf{x} \, \ln \mathcal{Z}(\alpha, \gamma_1, ..., \gamma_N) \tag{6.13}$$

for the energy constraint and

$$C_n = -\frac{\partial}{\partial \gamma_n} \int \mathrm{d}^2\mathbf{x} \, \ln \mathcal{Z}(\alpha, \gamma_1, ..., \gamma_N) \tag{6.14}$$

for each Casimir constraint. Here $\mathcal{Z}$ is the local normalization, or the partition function, of the probability distribution given as

$$\mathcal{Z} = \int \mathrm{d}\tilde{q} \, \exp\left(-\alpha \langle \psi \rangle \tilde{q} - \sum_{i=n}^{N} \gamma_n \tilde{q}^n \right), \tag{6.15}$$

thus, $\mathcal{Z}$ is constructed such that

$$\int \mathrm{d}\tilde{q} \, \rho(\tilde{q}|\mathbf{x}) = 1. \tag{6.16}$$

Determining the Lagrange multipliers for given values of N constraints is numerically difficult and its solution is not tackled in this study.

However, we can proceed by reducing the dimensionality of the problem. Ironically this is achieved by first considering the case of infinite dimensions. Constraining the entropy of the flow the first N polynomial Casimirs of the flow is a truncated version of the exact constraint. To constrain by *all* Casimirs of the flow we constrain by the global potential vorticity distribution discussed in Section 6.2. The constraint is given by

$$\Pi(\tilde{q}) = \int \mathrm{d}^2\mathbf{x} \, \rho(\tilde{q}|\mathbf{x}), \tag{6.17}$$

and the Lagrange multiplier becomes a function of $\tilde{q}$, $\gamma(\tilde{q})$. The corresponding variational problem

$$\frac{\delta S}{\delta \rho} + \alpha \frac{\delta}{\delta \rho} \left(-\frac{1}{2} \int \mathrm{d}^2\mathbf{x} \mathrm{d}\tilde{q} \, \langle \psi \rangle \tilde{q} \rho - E \right) - \frac{\delta}{\delta \rho} \left(\int \mathrm{d}^2\mathbf{x} \mathrm{d}\tilde{q} \, \gamma(\tilde{q}) \rho - \Pi(\tilde{q}) \right) = 0, \tag{6.18}$$

now gives the solution

$$\rho(\tilde{q}|\mathbf{x}) = \frac{1}{\mathcal{Z}} \exp\left(-\alpha\langle\psi\rangle\tilde{q} - \gamma(\tilde{q})\right), \tag{6.19}$$

for the probability distribution. Substituting (6.19) into (6.17) we obtain the expression

$$\Pi(\tilde{q}) = e^{-\gamma(\tilde{q})} \int \mathrm{d}^2\mathbf{x}\, \frac{e^{-\alpha\langle\psi\rangle\tilde{q}}}{\mathcal{Z}}. \tag{6.20}$$

The integral here is a function of potential vorticity only and we can write $\gamma(\tilde{q})$ in terms of Π and the integral. This allows us to eliminate the Lagrange multiplier corresponding to the Casimir constraint leaving us with only α to find. Eliminating γ from (6.19), we obtain the probability distribution in terms of the Lagrange multiplier, α, and the global distribution, Π:

$$\rho(\tilde{q}|\mathbf{x}) = \frac{1}{\mathcal{Z}(\mathbf{x}) \int \mathrm{d}^2\mathbf{x} \left(\frac{e^{-\alpha\langle\psi\rangle\tilde{q}}}{\mathcal{Z}(\mathbf{x})}\right)} \Pi(\tilde{q}) e^{-\alpha\langle\psi\rangle\tilde{q}}. \tag{6.21}$$

A numerical method can be constructed for the two-dimensional problem of optimizing the values of α and $\mathcal{Z}$ simultaneously thus finding the maximum entropy distribution from only knowledge of global quantities. The details of the numerical method is given in Appendix G and the results of this methodology follow.

6.5.4 Reconstruction of statistics

In this section we reconstruct the statistics from the maximum entropy distribution, (6.21), by optimizing for the Lagrange multiplier, α. The resulting values for α in simulations $\mathrm{FD}_1, ..., \mathrm{FD}_7$ are shown in Figure 6.7. We see that the value of α has a strong dependence on the strength of wind stress and time. Figure 6.7a shows the evolution of α during spin up, we see a strong reduction in the value of α at short times with the value of α fluctuating about a statistically steady state value for long

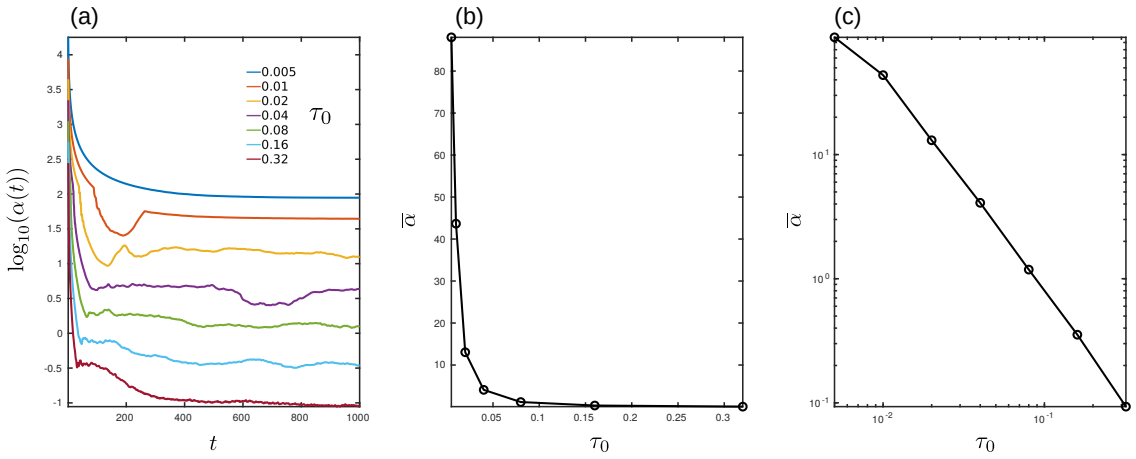


Figure 6.7: (a) Shows the evolution of $\log_{10}\alpha$ with time. We see a very quick relaxation to a steady state value with fluctuation around this value in steady state being small. (b) Shows the steady state mean value of α against the strength of the wind stress. We see at low forcing have a very high sensitivity of entropy to energy while at high forcing the sensitivity is low. (c) Shows mean value of α against wind stress strength, now on a log-log scale. We see a near power-law behaviour with an approximate slope of -1.7.

times. In steady state the fluctuation around the time-mean value is very small, this supports the assumption made in Section 6.5.2 to derive equation (6.12). Also shown is the dependence of the time-mean value of α on the wind stress strength in Figure 6.7b and c. The steady state sensitivity of entropy to energy is drastically decreased with the strength of wind stress suggesting that the entropy of the system becomes insensitive to perturbations in energy at high wind stress. Figure 6.7c displays a near power-law dependence of $\overline{\alpha}$ on τ_0 with an exponent of approximately -1.7, the explanation of this feature is left for future work. The increased insensitivity of entropy to energy variations at high wind stress has an interesting parallel to the concept of eddy saturation (e.g. Munday et al., 2013) which will be discussed briefly in Chapter 7 as an avenue for future consideration.

Interpretations of the Lagrange multiplier, α, come with a caveat; the numerically determined value of α is only as accurate as the maximum entropy hypothesis and, in particular, the mean field approximation for the energy. To test the accuracy of

these devices we compare the reconstructed statistics from the distribution, (6.21), and the diagnosed statistics from the numerical simulations.

Figure 6.8 shows diagnostics comparing the maximum entropy distribution evaluated using (6.21) with the statistics diagnosed from the numerical simulation itself, that is, the 'truth'. In short, despite displaying encouraging qualitative agreement with the simulations, the maximum entropy distribution does not fully capture the quantitative statistical details of simulations $\mathrm{FD}_1, ..., \mathrm{FD}_6$. Figure 6.8a shows a comparison between the reconstructed and diagnosed $\overline{\langle q \rangle} - \overline{\langle \psi \rangle}$ relation for simulation FD_6. The maximum entropy reconstruction shows good qualitative agreement but seems smoothed compared to the diagnosed relationship. Also for simulation FD_6, Figure 6.8b compares the $\overline{\mathrm{SD}(q)} - \overline{\langle \psi \rangle}$ relation and we see that the quantitative agreement is poorer. The standard deviation is underestimated in the centre of channel while being overestimated in the flanks. It must be noted that the maximum-entropy reconstruction of these statistical represents a drastic improvement over the energy-enstrophy constrained theory.

Figure 6.8c and d compare the diagnosed and reconstructed skewness and kurtosis for simulation FD_7. While giving qualitatively consistent features, the reconstructed skewness and kurtosis display significant deviations near to the boundary. In addition, Figure 6.8e and f compares the probability distribution, predicted and diagnosed, for simulations FD_4 and FD_5 at $y = 0$ and $y = 2.01$ respectively. In Figure 6.8e the trimodal nature of the distribution is successfully captured but we can see that the maximum entropy distribution overestimates the weight of the central peak compared to the side peaks. In Figure 6.8f, we again can see a good qualitative agreement but in the details there are deficiencies. Many of the ways in which the maximum entropy statistics deviates from the simulations can be interpreted as a degradation of the strong persistent mixing barrier which the simulations $\mathrm{FD}_1, ..., \mathrm{FD}_6$ exhibit. The statistical nature and consequences of this mixing barrier has been extensively

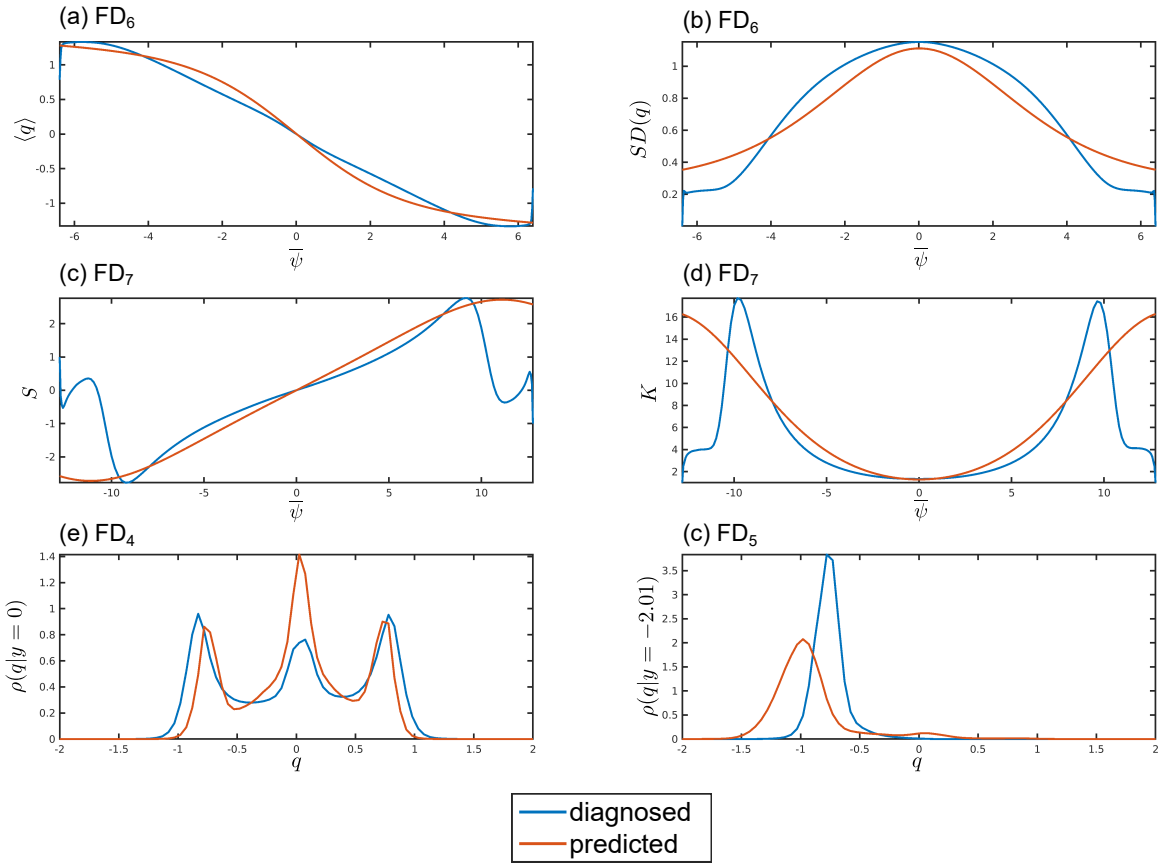


Figure 6.8: Diagnostics testing reconstructed maximum entropy statistics. (a) Comparison of the time-averaged mean potential vorticity against time-averaged streamfunction for simulation FD_6. (b) Comparison of the time-averaged standard deviation of potential vorticity against time-averaged streamfunction for simulation FD_6. (c) Comparison of the time-averaged skewness of potential vorticity against time-averaged streamfunction for simulation FD_7. (d) Comparison of the time-averaged kurtosis of potential vorticity against time-averaged streamfunction for simulation FD_7. (e) Comparison of the maximum entropy probability distribution and the diagnosed distribution from simulation FD_4, distributions are evaluated for the centre of the channel, $y = 0$.(f) Comparison of the maximum entropy probability distribution and the diagnosed distribution from simulation FD_5, distributions are evaluated for the centre of the channel, $y = 0$.

studied, for this model, in David et al. (2017).

In summary, the maximum entropy distribution, (6.21), produces an encouraging qualitative reproduction of the flow statistics. However, the quantitative reproduction is still lacking despite the suggestive evidence for the entropy being maximized presented in Section 6.5.2. The derivation of the maximum entropy distribution relies on two main assumptions in the equilibrium, Miller-Robert-Sommeria, statistical mechanics: a) maximization of entropy; and b) the mean field approximation. We argue that the deficiencies of the maximum entropy distribution presented in this numerical experiment can be attributed to the break-down of the mean field approximation in a forced dissipative system. This will be discussed in detail in the following section.

6.6 The mean-field approximation

A possible reason for the lack of success of the predicted probability distribution function, (6.21), is that the maximum entropy principle is not at work. Nevertheless, we believe that the analysis presented in Section 6.5.2 provides sufficient evidence to look for other reasons why (6.21) fails to quantitatively capture the statistics. In this section we consider the mean field approximation of Miller-Robert-Sommeria equilibrium statistical mechanics. The energy of the flow is given as

$$E[q] = -\frac{1}{2}\int \mathrm{d}\mathbf{x}\,\psi(\mathbf{x})(q(\mathbf{x}) - \beta y), \tag{6.22}$$

because of the presence of ψ it is not clear how make the necessary substitution to write E as a functional of ρ allowing us to tackle this constraint analytically. We rewrite the energy as

$$E[q] = -\frac{1}{2}\int\int \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'\,(q(\mathbf{x}) - \beta y)G(\mathbf{x},\mathbf{x}')q(\mathbf{x}') \tag{6.23}$$

where $G(\mathbf{x}, \mathbf{x}')$ is the Green's function of the differential operator defined by, $q = \nabla^2\psi + \beta y$. Now, by swapping the potential vorticity field for its average value, and defining $\langle\tilde{q}\rangle \equiv \nabla^2\langle\psi\rangle + \beta y$, we get;

$$E_M[\rho] = -\frac{1}{2}\int \mathrm{d}\mathbf{x}\,\langle\psi\rangle(\langle\tilde{q}\rangle - \beta y). \tag{6.24}$$

This step is the *mean-field approximation*, often used in models of condensed matter physics. In essence, we are saying that, rather than considering the interaction energy between all pairs of potential vorticity patches, we consider that each patch of potential vorticity only feels mean effect of all other patches.

For an ideal fluid in statistical equilibrium, the mean-field approximation ceases to be approximate and becomes exact due to the non-local aspect of the Green's function, G (Bouchet and Venaille, 2012); for example, two-dimensional vortex dynamics had a logarithmic Green's function leading to long-range vortex-vortex interactions. The mean-field formulation, however, necessitates two related properties of ideal equilibria: a) that the mean eddy potential vorticity flux is zero, $\nabla \cdot \langle\mathbf{u}'q'\rangle = 0$, which is a consequence of the equilibrium-state q-ψ relation, $\langle q\rangle = f(\langle\psi\rangle)$; and b) that neighbouring macro-cells of the flow are uncorrelated. These properties are fundamentally linked with the mean field approximation and are manifestly not satisfied in a forced-dissipative statistically steady state. Therefore, we suggest that while the maximization of entropy might be an useful organizing principle in forced-dissipative flow, the mean-field approximation remains only a crude approximation. We propose two potential avenues for future study to tackle this problem.

- It may be possible to find a coarse-graining (i.e., macro-cells) which is partially Lagrangian to reduce difference between $E[q(\mathbf{x})]$ and $E_M[\rho(\tilde{q}|\mathbf{x})]$. This is the approach taken in Jung et al. (2006) which produces good agreement between experiment and equilibrium theory by moving into a frame of reference moving

at the phase speed of a large scale Rossby wave. However, this is unlikely to work in the presence of multiple wave modes as is the case with simulations $\mathrm{FD}_1, ..., \mathrm{FD}_6$, see discussion in David et al. (2017).

- Alternatively, we speculate that a perturbation method could be applied to the mean-field approximation to yield a more realistic probability distribution. However, precisely how this might be achieved remains to be determined and is subject of future work.

Indeed, it may be that a combination of the two approaches will provide the means of deriving forced-dissipative statistics from the maximum entropy principle.

6.7 Concluding remarks

In this study we have shown how an eddy mixing entropy can be used as a measure of turbulent disorder. By deriving the influence of forcing and linear drag, we were able to use entropy to describe the turbulence in a freely-decaying and forced-dissipative flow. The evolution of entropy describes the three stages of the eddy life cycle and eddy-mean interaction: growth of instability, formation of large scale coherent structures and steady state fluctuations. In particular, the eddy production of entropy, which has been focus of much theoretic inquiry, can be explicitly computed from data. This will inform work on stochastic parametrization by describing the disorder in a turbulent jet in a way that links to both statistical physics and information theory.

The relationship between the temporal evolution of entropy and the maximum entropy principle was considered in Section 6.5. Under the assumption of maximum entropy it was found that the time evolution of entropy was set by the time evolution of its constraints. Suggestive evidence was found that the entropy is maximized in the model simulations considered in this study. It is clear that if a variational problem can be used to infer the statistics then the number of Casimir constraints has to be

large.

With this evidence for the maximum entropy principle being a physically meaningful candidate for describing the behaviour of turbulence in the system studied here, we considered the problem of inferring the sub-grid scale statistics. This is equivalent to inferring the Lagrange multipliers, used in the maximum entropy variational problem, from the constraints applied. We presented the mathematical formulation of this problem in Section 6.5 and showed how the dimensionality could be reduced given knowledge of the global potential vorticity distribution. Further, we reconstructed the maximum entropy statistics from knowledge of the energy, global potential vorticity distribution, and zonal mean streamfunction as a functions of time. We find that although the maximum entropy statistics reproduces qualitatively representative features of the flow, quantitative agreement is an improvement on the energy-enstrophy theory, but is not yet in complete agreement. In Section 6.6, the mean-field approximation was discussed as a potential culprit for the quantitative disagreement and avenues for future investigation were proposed.

In this chapter we have presented eddy-mixing entropy as both a descriptive tool and a dynamically balanced quantity in a barotropic turbulent jet. We have also demonstrated the relationship between the statistical mechanics of forced-dissipative flow and well-known globally balanced quantities such as the energy and enstrophy of the flow. In doing so we were able to provide evidence for the action of the maximum entropy principle at work in a forced-dissipative system. The question of the usefulness of statistical mechanics theories, such as the Miller-Robert-Sommeria theory, in understanding the statistically steady states of two-dimensional and geophysical turbulence has received much attention (e.g. Wang and Vallis, 1994; Jung et al., 2006; Qi and Marston, 2014; Dritschel et al., 2015) but remains somewhat unclear. By explicitly considering the evolution of eddy mixing entropy in a forced-dissipative model we are able to demonstrate the importance and utility of eddy mixing entropy in the

study of forced-dissipative geophysical turbulence, opening the door to revisiting the application of statistical mechanics to ocean mesoscale eddy parameterizations (cf. Kazantsev et al., 1998).

Chapter 7

Conclusion

7.1 Summary of results

Here we summarize the key results of this thesis, presented in Chapters 4, 5, and 6.

In Chapter 4, a study of potential vorticity statistics in a barotropic channel ocean was presented. It was found that jets concomitant with strong mixing barriers exhibit multi-modal and mixed probability distributions with self-similar consequences for low-order statistical moments over a wide range of forcing strength. The characteristic skewness/kurtosis behaviour can understood when the large-scale motion of the mixing barrier is modeled as a Bernoulli process. When the influence of the mixing barrier is eliminated, either by exploiting dynamics (jets with no mixing barrier behaviour) or diagnostically (streamfunction following frame of reference), the characteristic multi-modal nature of the statistics can be removed. In agreement with previous studies (e.g. Hughes et al., 2010), this points to the utility of skewness and kurtosis in identifying ocean jets as well in providing an important characterization of jet and mixing barrier statistics in an idealized context. In particular, an important relationship is found between mixing barriers in the dynamics and mixed probability distributions in the statistics.

In Chapter 5, the results of Chapter 4 were considered in the context of a realistic, observation constrained, ocean model, SOSE. It was found that potential vorticity skewness and kurtosis can be used to identify potential vorticity fronts at depth, although, the identified fronts could not always be associated with jets and mixing barriers simultaneously. By considering statistics of potential vorticity in different frames of reference, three categories of potential vorticity front were identified: 1) potential vorticity fronts associated with the topographically localized features of the Agulhas retroflection and Malvinas extension; 2) potential vorticity associated with the boundaries between the core of the Antarctic Circumpolar Current and mode water to the north; and 3) potential vorticity fronts formed by the entrainment of high potential vorticity at the base of the deep winter mixed layer. Evidence is presented (Osborne-Cox diffusivity and statistical sub-populations) that these latter fronts act as partial barriers to mixing along neutral density surfaces, becoming more impermeable on denser surfaces.

In Chapter 6, a study was presented of the applicability of eddy-mixing entropy and the maximum entropy principle in a forced-dissipative barotropic channel ocean. It was found that the eddy-mixing entropy acts as a dynamically balanced quantity with sources provided by the eddies, via linear instability, and sinks provided by both the eddies, via the emergence of large-scale structures, and the constant effect of linear drag. Further, evidence was found that eddies act to maximize the entropy despite the forced-dissipative nature of the flow, potentially providing a useful means of reconciling equilibrium statistical mechanics theory with more realistic non-equilibrium turbulent flow. This provides an important first step towards an ocean mesoscale parameterization based on a statistical mechanics theory of turbulence.

In this concluding chapter we will consider in detail two common themes which have formed part of the discussion of the three chapters summarized above (Sections 7.2 and 7.3). In Section 7.4, some additional considerations for a statistical mechanics

theory of the ocean mesoscale are given. In Section 7.5, some final remarks are given.

7.2 The importance of Lagrangian information

The importance of Lagrangian information is a common theme that has emerged throughout this thesis, specifically in Chapter 4 (streamfunction following statistics, Section 4.6), Chapter 5 (neutral density following statistics, Section 5.3.3), and Chapter 6 (instantaneous streamfunction contours as macro-cells, Section 6.5.2). Discussed in this section are some further considerations for how Lagrangian or quasi-Lagrangian statistics are important for understanding ocean mesoscale statistics.

7.2.1 Galilean transformed mean and Rossby wave spectra

In Chapter 4, a Galilean transformed mean methodology is used in an attempt to remove the multi-modal nature of the statistics along streamfunction contours. In addition to the results presented in Chapter 4, Rossby wave spectra were found, given in Appendix B (Figures B.1, B.2 and B.3), which display some interesting features of the Rossby wave dynamics. We highlight these features here.

- The fact that we see distinct peaks in the spectra is a consequence of the necessary quantization of the waves due to the period boundary conditions.

- By considering secondary peaks in the spectra it is found that the underlying dispersion relation is multi-valued with Rossby waves with the same wavenumber leading to multiple wave speeds, pointing to a complex interaction between waves and time-varying background flow.

- As can be seen comparing Figures B.1, B.2 and B.3 the structure of the spectra changes radically as the wind stress strength is varied, in contrast to the high

degree of self-similarity observed in the spatial statistics presented in Chapter 4.

Understanding these Rossby spectra in their own right offers an interesting avenue for future work on wave-mean flow interactions but also may provide a means of deriving a time-averaging method which incorporates large-scale flow structures by decomposing the flow onto a set of the most influential propagating Rossby waves.

7.2.2 Topology and coherent structures

In the previous section, and in Chapter 4, the influence and characteristics of propagating Rossby waves were considered. Also of importance are other typical coherent structures of geostrophic turbulence such as coherent eddies. The effect of these coherent eddies in the form of Agulhas rings was seen in Chapter 5 where the drifting ocean rings have a characteristic signal in simple statistics of the Ertel potential vorticity.

The simulations of the barotropic model used in this thesis also exhibit flow regimes which contain coherent eddies, for example, see Figure 4.4. Indeed, when the analysis of Chapter 6, Sections 6.5.3-6.5.4, is applied to flows containing a coherent eddy, the optimization for the Lagrange multiplier fails. This is because the mean-field energy approximation is not appropriate when defined with a zonal mean and applied to a flow containing with topologically distinct coherent eddies which contain much of the energy. In a recent study, Frishman et al. (2017), highlight the importance of the topological distinction between jets and propagating coherent eddies; while both lead to zonally symmetric long-time average flow, the latter cannot be explained using commonly used theoretical arguments, including the maximum entropy principle.

A possible means of consistently treating coherent eddies within a statistical mechanics framework might come from the methodology of Stanley (2017). When

topologically distinct regions of an ocean neutral surface, defined by closed pressure contours, are identified an exact geostrophic streamfunction can be found. This methodology could be extended to identify topologically distinct regions which are characterized by a potential vorticity/streamfunction relation suggestive of the result of Miller-Robert-Sommeria (see Section 3.3.3). The methodology of Stanley (2017) is also an attractive avenue for empirically analyzing coherent eddies in simple models as well as extending the statistical mechanics formalism to realistic ocean models.

7.3 The importance of higher order Casimirs

The study of two-dimensional and geostrophic turbulence has been dominated by the idea that there are two invariants of motion which need to be considered: *energy* and *enstrophy*. This idea begins with the formulation of Kraichnan theory (Kraichnan, 1967, 1971; Leith, 1968; Batchelor, 1969) and has been used to formulate equilibrium statistical mechanics-like theories (Salmon et al., 1976; Bretherton and Haidvogel, 1976; Jung et al., 2006). The reasoning for ignoring the infinitely many other invariants of motion lies in the *robustness* of quadratic invariants (energy and enstrophy): the non-linear term in the equations of motion continues to conserve quadratic invariants under spatial or Fourier truncation.

Other studies show how higher order Casimirs are important to describe flow statistics (Abramov and Majda, 2003) as well as argue that these invariants exhibit turbulent cascades (Polyakov, 1993; Falkovich and Hanany, 1993; Eyink, 1996) although the question of whether the cascades are inverse or direct is unresolved. More recently, Bowman (2013) was able to find evidence for a direct cascade of the fourth order Casimir, C_4, in a sufficiently high resolution simulation of two-dimensional turbulence.

In this thesis evidence is found that supports the notion that it is important to

consider the behaviour of higher order Casimirs, especially in forced-dissipative flow. In forced-dissipative flow, the robustness of Casimirs (or a lack thereof) does not matter. We can see this by considering the evolution of the global distribution of potential vorticity, $\Pi(q)$, the conservation of which is equivalent to conserving all of the infinite number of Casimirs. By integrating the tendency equation for ρ, (6.4), in Chapter 6 the tendency of Π is

$$\frac{\partial \Pi}{\partial t} = \int \mathrm{d}^2\mathbf{x} \left(\frac{\partial}{\partial \tilde{q}} \left(r(\tilde{q} - \beta y)\rho \right) \right) - \int \mathrm{d}^2\mathbf{x} \left(\frac{\partial \tau}{\partial y} \frac{\partial \rho}{\partial \tilde{q}} \right) \tag{7.1}$$

where all notation is as defined in the previous chapters. Note that, with typical boundary conditions, the advective term does not have an explicit effect on the evolution of Π, that is, probability cannot be fluxed out of the domain. It is clear that in general one cannot infer that higher order Casimirs, which are the raw moments of Π defined as

$$C_n = \int \mathrm{d}q \, q^n \Pi(q), \tag{7.2}$$

become unimportant. This is because the input of each moment of equation (7.1) will depend on the particular spatial structure of the forcing, τ, even in a truncated numerical model. With the addition of the drag term we can see how any moment can be maintained in a dynamical balance and will be relevant for the turbulent flow statistics.

In Chapter 4 we found evidence of this behaviour in multi-modal global probability distributions. While not shown explicitly their multi-modal nature can be inferred by inspection of Figure 4.14. These require a high number of Casimirs (raw moments) to be described, see discussion in Section 4.7.2. Further, the evidence of statistical sub-populations within the potential vorticity statistics on a neutral density surface in Chapter 5, Section 5.5 again suggests the need for a large number of Casimirs to describe the dynamically relevant global probability distribution function.

Additionally, important evidence of the relevance of high order Casimirs is presented in Chapter 6 in an entropic framework. It is found that a large number of Casimir constraints are needed to the temporal evolution of entropy, under the assumption of entropy maximization, in a forced-dissipative flow. Again this can be related to the multi-modal structure of the global potential vorticity distribution, which is the dynamical quantity which constrains the entropy maximization in Miller-Robert-Sommeria statistical mechanics.

Thus, one of the conclusions of this thesis is that high order Casimirs are *essential* when constructing a statistical theory of the ocean mesoscale characterized by multi-modal distributions of potential vorticity as a consequence of potential vorticity fronts, jets, mixing barriers and coherent eddies.

7.4 Towards a statistical theory of the ocean mesoscale

7.4.1 Entropy production

One of the ways that has been suggested to extend equilibrium statistical mechanics to forced-dissipative geophysical flow is via the Maximum Entropy *Production* Principle (MEPP). Kazantsev et al. (1998) propose the use of a MEPP methodology to derive a sub-grid parametrization for barotropic geophysical flows. Here we will discuss how the results presented in Chapter 6 deviate from the proposal of Kazantsev et al. (1998).

In Chapter 6 an equation for the entropy tendency was derived, equation (6.6). Based on the empirical evidence that entropy fluctuated around a steady time-mean value at long time we can write that

$$\overline{P} = r \tag{7.3}$$

in a forced-dissipative statistically steady state, where the over-line denotes a time-mean. This implies that the time-mean *advective production of entropy* is completely determined by the linear drag coefficient. This reduces the usefulness of the MEPP in a forced-dissipative flow as the steady state advective production of entropy behaves in such a way to balance its sinks (and sources see Appendix D). An interesting extension to the work presented here would be to consider the situation where there is an additional source of entropy via the addition of stochastic noise.

The relationship given in (7.3) suggests an interesting parallel with the recent model of eddy saturation Marshall et al. (2017) where it was found that, in an idealized Southern Ocean, the volume transport scales with the linear drag and the Eady growth rate becomes directly proportional to the linear drag coefficient under these conditions. Curiously, equation (7.3) shows that the advective production of entropy is set by the linear drag coefficient, prompting the questions: 1) Is the advective production of entropy proportional to the barotropic linear shear instability growth rate in the model considered in this thesis? This can be approached in two ways; analytically by considering the evolution of entropy in the linearized equations of motion and numerically, by solving for the linear modes for the background states evaluated from the non-linear model. 2) Would the advective production of entropy be proportional to the Eady growth rate if the analysis of Chapter 6 were applied to a two-layer quasi-geostrophic model? To begin with, it would be essential to repeat much of the analysis of Chapter 6 in the case of a two-layer model before considering the relationship to the Eady growth rate. These questions are left for future consideration.

7.4.2 Extension to more realistic flows

A key question which arises from the work presented in this thesis and in particular Chapter 6 is: how do statistical mechanics theories translate to the real ocean?

The statistical mechanics theory presented here has been extended to: the three-dimensional axisymmetric Euler equations (Naso et al., 2010; Thalabard et al., 2014, 2015); shallow water equations (Warn, 1986; Weichman and Petrich, 2001; Renaud et al., 2016); as well as the Boussinesq equations for stratified turbulence (Venaille et al., 2017). All these studies, however, retain the inherent deficiency of assuming *ideal* flow. The results presented in Chapter 6 gives a preliminary route, for the first time, to considering the case of forced-dissipative flow.

In order to extend the results of Chapter 6 some important questions must be addressed. For example, to what extent is the system in quasi-equilibrium (Section 6.5.2)? Can the approximate quasi-equilibrium behaviour be seen in different flows (e.g. plumes in stratified turbulence)? An obvious framework in which to start to answer these questions is that of baroclinic quasi-geostropic dynamics as it contains an internal deformation radius which is crucial for representing baroclinic geostrophic eddies.

7.5 Concluding remarks

Ocean mesoscale turbulence is highly inhomogeneous, anisotropic, forced, dissipative, exhibits non-Gaussian statistics and is characterized by non-linear interactions across disparate length-scales. The traditional homogeneous turbulence paradigm is unsuitable for describing large-scale ocean turbulence as it assumes homogeneity, isotropy, an inertial range and a turbulent cascade. Miller-Robert-Sommeria statistical mechanics, on the other hand, does not make these assumption apart from the prohibitive assumption of no forcing and no dissipation.

This thesis combines empirical analysis, obtained from simplified numerical experiments and a realistic ocean model, with theoretical considerations in order to identify the essential properties a theory of the ocean mesoscale must have. In ad-

dition to the particular results presented in each of Chapters 4, 5 and 6 evidence is presented for two important factors that should be considered when describing the ocean mesoscale: 1) Lagrangian or quasi-Lagrangian frames of reference are key to distinguishing large-scale coherent structures (mixing barriers, coherent eddies) from the smaller scales which act more stochastically; and 2) higher order Casimirs are necessary to describe the multi-modal potential vorticity distributions, arising from coherent structures, as well as playing an important role in the maximum entropy principle.

The ocean mesoscale is an integral component of the climate system as a whole. However, its numerical representation lies beyond computational power for the foreseeable future, making a more sophisticated theoretical description a pressing concern for science. The results presented in this thesis represent a contribution toward the development of such a theoretical understanding of the ocean mesoscale.

Appendix A

Boundary condition for finite deformation radius

McWilliams (1977) argues that if we are to derive the quasi-geostrophic equations via a small parameter expansion of the primitive equations then we had better ensure that the boundary conditions are consistent under the same expansion. In simply-connected domains the complications arising from this are trivially satisfied. However in multiply connected domains additional mass and momentum constrains must be satisfied as an additional prognostic equation. The description presented here follows from the work of Nadeau (personal communication).

The first step in applying boundary condition (2.27) is to realize that we can write our streamfunction as

$$\psi = \tilde{\psi} + \frac{\Gamma(t)h(y)}{2}. \tag{A.1}$$

$\tilde{\psi}$ is the solution to the inhomogeneous equation

$$\nabla^2\tilde{\psi} - \frac{\tilde{\psi}}{L_D^2} = \zeta \tag{A.2}$$

with Dirichlet boundary conditions of

$$\tilde{\psi}|_{N,S} = 0. \tag{A.3}$$

$\zeta = q - \beta y$ is the relative vorticity. $h(y)$ is the solution to the homogeneous equation

$$\nabla^2 h - \frac{h}{L_D^2} = 0 \tag{A.4}$$

and satisfies the Dirichlet boundary condition

$$h|_{N,S} = \pm 1, \tag{A.5}$$

and can be precomputed at model initialization.

The shallow water quasi-geostrophic momentum equations can be written as

$$\frac{\partial u_g}{\partial t} + \mathbf{u_g} \cdot \nabla u_g - f_0 v_a - \beta y v_g = -g\frac{\partial \eta_a}{\partial x} - \kappa u_g - \nu_h \nabla^4 u_g + \tau \tag{A.6}$$

for the zonal direction, and

$$\frac{\partial v_g}{\partial t} + \mathbf{u_g} \cdot \nabla v_g + f_0 u_a + \beta y u_g = -g\frac{\partial \eta_a}{\partial y} - \kappa v_g - \nu_h \nabla^4 v_g \tag{A.7}$$

for the meridional direction. The velocity and surface elevations have been decomposed into a part which is in geostrophic balance, denoted by the g subscript, and a part representing the small ageostrophic deviations, denoted by a subscript. Let us consider the balance of zonal momentum along a latitude band, l. We get the integral balance

$$\oint_l \mathrm{d}x \left[\frac{\partial u_g}{\partial t} + \mathbf{u_g} \cdot \nabla u_g - f_0 v_a - \beta y v_g + g\frac{\partial \eta_a}{\partial x} + \kappa u_g + \nu_h \nabla^4 u_g - \tau \right] = 0. \tag{A.8}$$

Taking the integral and applying appropriate boundary conditions many terms drop out and we get

$$\oint_l \mathrm{d}x \left[\frac{\partial u_g}{\partial t} + \mathbf{u_g} \cdot \nabla u_g - f_0 v_a + \kappa u_g + \nu_h \nabla^4 u_g - \tau \right] = 0. \tag{A.9}$$

The term which gives the trouble here is is the $f_0 v_a$ term which is not easily calculated in a potential vorticity model. To access this term we need to integrate the mass balance equation over the area under the l. Doing this yields the integral balance

$$\frac{D}{Dt} \iint_A \mathrm{d}A\, \eta = \oint_l \mathrm{d}x \left[H v_a + v_g \eta \right], \tag{A.10}$$

where we define the surface height is given by $h = H + \eta$. Rearranging, multiplying by f_0 and using the definition of the streamfunction in this model

$$\psi = \frac{g\eta}{f_0^2}, \tag{A.11}$$

we get the expression

$$\oint_l \mathrm{d}x\, f_0 v_a = \frac{D}{Dt} \iint_A \mathrm{d}A \frac{\psi}{L_D^2} - \oint_l \mathrm{d}x \frac{\partial \psi}{\partial x} \psi. \tag{A.12}$$

Substituting this back into (A.9) we get a closed evolution equation

$$\frac{D}{Dt} \left[\oint_l \mathrm{d}x\, \frac{\partial \psi}{\partial y} + \iint_A \mathrm{d}A\, \frac{\psi}{L_D^2} \right] = \oint_l \mathrm{d}x \left[\kappa \frac{\partial \psi}{\partial y} + \nu_h \frac{\partial^5 \psi}{\partial y^5} \right] \tag{A.13}$$

for the momentum balance along l. By substituting the decomposition, (A.1), into this equation we can solve for the boundary condition at each time-step thus ensuring momentum conservation for all times in the model.

Appendix B

Galilean transformed mean and Rossby wave spectra

Here, we examine, in more detail, the Galilean transformed mean employed in the statistical analysis of Chapter 4.

A key step in the analysis of Jung et al. (2006) is to move into a frame of reference moving at the phase speed of a large-scale barotropic Rossby wave which dominates the flow via a Galilean coordinate transformation defined by

$$u' = u - v_{ph}, \tag{B.1}$$

$$x' = x - v_{ph}t, \tag{B.2}$$

$$y' = y, \tag{B.3}$$

$$t' = t. \tag{B.4}$$

Under Galilean transformations, potential vorticity transforms like a scalar, that is

$$q' = q, \tag{B.5}$$

while the streamfunction, which is not a scalar, depends on the velocities which are

now reduced by v_{ph}, giving

$$\psi' = \psi + \frac{v_{ph}}{2} y, \tag{B.6}$$

as the correct streamfunction in the new frame of reference. In moving into a frame of reference with speed v_{ph}, the Rossby wave is 'frozen out' when the Eulerian time-average is taken. All propagating Rossby waves with non-zero phase speed in a given frame of reference will always average out, giving a zonally symmetric time-average with a sufficiently long time average.

In Chapter 4, the analysis of Jung et al. (2006) was repeated but failed to remove the multi-modality from the statistics, associated with the large-scale meandering of the mixing barrier due to Rossby waves. This failure can be explained by the complex behaviour of Rossby waves in the simulations used in Chapter 4, manifested in two ways.

Firstly, we note that the Rossby wave dispersion relation with uniform background flow is insufficient to describe the observed wave propagation. In this simple case the dispersion is

$$\omega = Uk - \frac{\beta k}{k^2 + l^2}, \tag{B.7}$$

where U is a background zonal velocity, k is the zonal wavenumber, l is the meridional wavenumber. This gives a corresponding phase speed of

$$v_{\mathrm{ph}} = U - \frac{\beta}{k^2}, \tag{B.8}$$

in the case of a zonally propagating Rossby wave ($l = 0$). When the flow is inspected to determine the wavenumber, k, and we compute the phase speed, we find that the latter is radically different from the observed phase speed. The assumption of uniform background flow is too restrictive to obtain an accurate dispersion relation for the simulated Rossby waves.

Secondly, by empirically searching for the phase-speed of the dominant Rossby

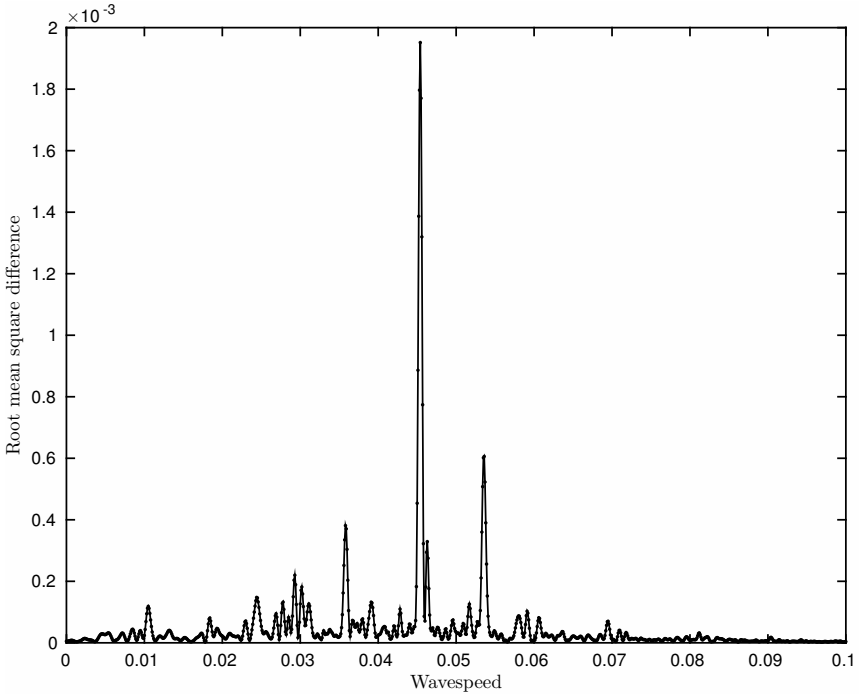


Figure B.1: Spectrum of root mean square difference between time mean and transformed time mean against test wavespeed for run A. We see peaks due to higher and lower wavenumber Rossby waves to either side of the dominant peak.

wave, an intricate spectrum of Rossby waves is uncovered. The method selects a range of Galilean transformations and measures the deviation of the resulting time-average potential vorticity from the zonally symmetric case; examples of the resulting Rossby wave spectra are shown in Figures B.1, B.2 and B.3. In Chapter 4, the largest peak is used to determine the dominant Rossby wave speed giving the Galilean transformed mean used in the analysis.

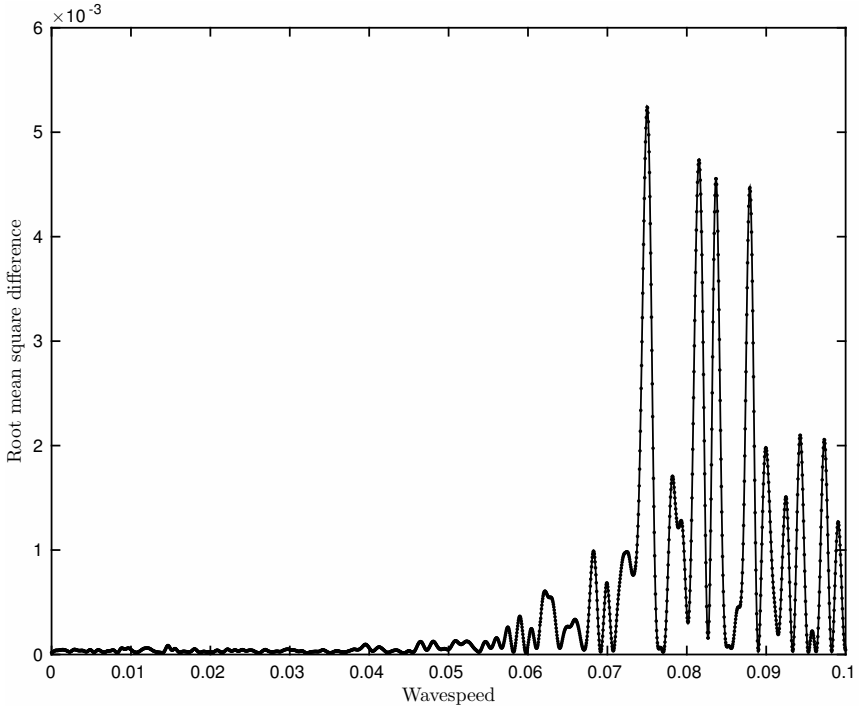


Figure B.2: Spectrum of root mean square difference between time mean and transformed time mean against test wavespeed for run D. This is one of the messiest of the spectra with large peaks near to the most dominant.

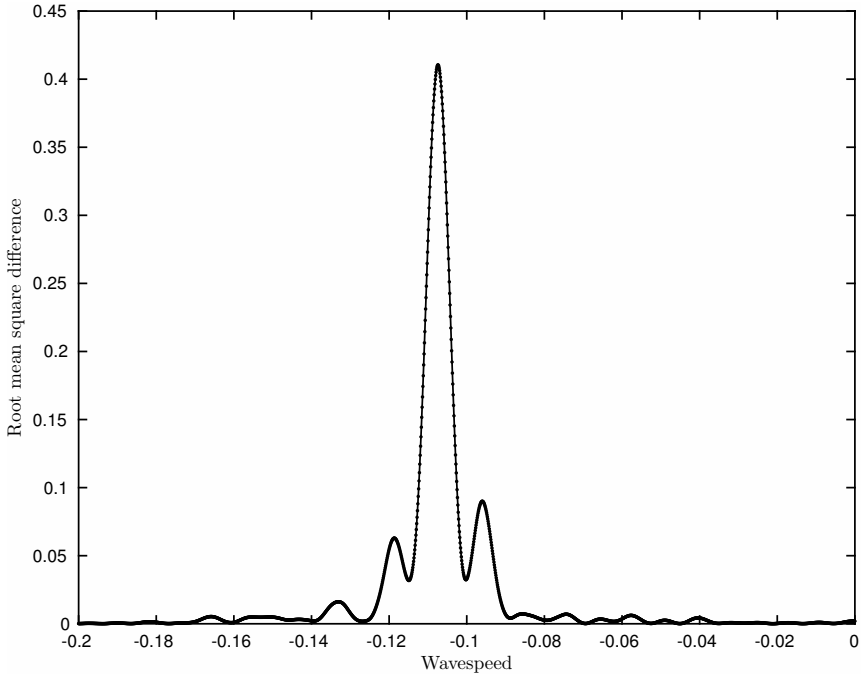


Figure B.3: Spectrum of root mean square difference between time mean and transformed time mean against test wavespeed for run G. For lower wavenumbers the peaks have become broader and it becomes consequently easier to find the dominant Rossby wave.

Appendix C

Numerical computation of entropy

It is important to note the difference between the discrete, or Shannon, entropy

$$S = -\sum_i \rho_i \ln \rho_i, \tag{C.1}$$

and the continuous, or differential, entropy we use in this study

$$S = -\int \mathrm{d}x\, \rho(x) \ln \rho(x). \tag{C.2}$$

One of the clear differences between these two entropies is that the continuous entropy can become negative whereas the discrete entropy is never less than zero. On the other hand the continuous entropy can be negative and indeed tends to negative infinity for the asymptotic limit of a delta-function. This means that we need to be careful when numerically evaluating an estimator for the continuous entropy. Naïvely using the standard discrete approximation for the integral in (C.2) leads to calculating a quantity proportional to the discrete entropy, (C.1). To find an approximation for, (C.2), we must evaluate the quantity

$$S \approx -\Delta x \sum_i \rho_i \ln \rho_i + \ln \Delta x, \tag{C.3}$$

where ρ_i becomes a histogram approximation to the distribution $\rho(x)$. However this method of approximation was found to be biased and introduced a systematic error into the results presented in this study.

Instead, we used a sample-spacing estimator for the distribution leading to an improved numerical approximation for the continuous entropy (Vasicek, 1976; Beirlant et al., 1997). The sample-spacing estimator relies of the idea that when the data is ordered, from smallest to largest value, then the reciprocal of the difference between two samples, separated by m spaces, is an estimator for the probability density. Using this, the following expression is found for the entropy

$$S \approx \frac{1}{N-m} \sum_{i=1}^{N-m} \ln\left(\frac{N+1}{m}(x^{(i+m)} - x^{(i)})\right), \tag{C.4}$$

where, N is the number of samples; m is the spacing size; and $x^{(i)}$ represents the i^{th} ordered sample. This method was used to evaluate the entropy throughout this study and is found to be considerably better than more naïve methods. The simpler methods proved to have a strong dependence on the choice of histogram bin width rendering them unusable for quantitative comparison with theory.

Appendix D

Effect of stochastic white noise on eddy-mixing entropy

Ignoring the non-local and non-linear terms but now including a white Gaussian noise, η, the potential vorticity equation

$$\frac{\partial q}{\partial t} = -r(q - \beta y) + g(y) + \eta, \tag{D.1}$$

represents an Ornstein–Uhlenbeck process (e.g. Gardiner, 2004) and the corresponding probability distribution satisfies the Fokker-Planck equation,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \tilde{q}}[r(\tilde{q} - \beta y)\rho] - g(y)\frac{\partial \rho}{\partial \tilde{q}} + \frac{\langle \eta^2 \rangle}{2}\frac{\partial^2 \rho}{\partial \tilde{q}^2}, \tag{D.2}$$

where $\langle \eta^2 \rangle$ is the variance of the noise.

We can derive the entropy tendency by substituting equation (D.2) into

$$\begin{aligned}
\frac{dS}{dt} &= -\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\rho\ln\rho, && \text{(D.3)}\\
&= -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left[\frac{\partial\rho}{\partial t}\ln\rho+\frac{\partial\rho}{\partial t}\right], && \text{(D.4)}\\
&= -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\ln\rho-\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\rho, && \text{(D.5)}\\
&= -\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\ln\rho, && \text{(D.6)}
\end{aligned}$$

where the final step comes from the normalization condition on the distribution, ρ. We rewrite (D.2) as

$$\frac{\partial\rho}{\partial t}=r\frac{\partial(\tilde{q}\rho)}{\partial\tilde{q}}-h(y)\frac{\partial\rho}{\partial\tilde{q}}+\frac{\langle\eta^2\rangle}{2}\frac{\partial^2\rho}{\partial\tilde{q}^2}, \tag{D.7}$$

where

$$h(y)=r\beta y+g(y). \tag{D.8}$$

Substituting, we obtain

$$\dot{S}=-\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left(\left(-h(y)\frac{\partial\rho}{\partial\tilde{q}}+r\frac{\partial(\tilde{q}\rho)}{\partial\tilde{q}}+\frac{\langle\eta^2\rangle}{2}\frac{\partial^2\rho}{\partial\tilde{q}^2}\right)\ln\rho\right). \tag{D.9}$$

We take each term separately in order to simplify the above expression. The first term is simplified, integrating by parts, as

$$\begin{aligned}
\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left(-h(y)\frac{\partial\rho}{\partial\tilde{q}}\ln\rho\right) &= \int \mathrm{d}^2\mathbf{x}\,h(y)\int \mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial\tilde{q}}\ln\rho, && \text{(D.10)}\\
&= -\int \mathrm{d}^2\mathbf{x}\,\left(h(y)\left([\rho\ln\rho]_{-\infty}^{+\infty}-\int \mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial\tilde{q}}\right)\right) && \text{(D.11)}\\
&= 0. && \text{(D.12)}
\end{aligned}$$

We can see that the integral above is zero from the fact that both ρ and $\rho\ln\rho$ vanish

at $\pm\infty$. The second term can also be simplified, integrating by parts twice, as

$$\begin{aligned}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\left(r\frac{\partial(\tilde{q}\rho)}{\partial\tilde{q}}\ln\rho\right) &= r\int \mathrm{d}^2\mathbf{x}\,\left([\tilde{q}\rho\ln\rho]_{-\infty}^{+\infty}-\int \mathrm{d}\tilde{q}\,\tilde{q}\frac{\partial\rho}{\partial\tilde{q}}\right) &\text{(D.13)}\\ &= -r\int \mathrm{d}^2\mathbf{x}\,\left([\tilde{q}\rho]_{-\infty}^{+\infty}-\int \mathrm{d}\tilde{q}\,\rho\right) &\text{(D.14)}\\ &= Ar, &\text{(D.15)}\end{aligned}$$

where A is the area of the domain. Here, we have, in addition to the appropriate boundary conditions, used the normalization condition for ρ. The third and final term, from the stochastic noise, cannot be simplified further. Thus, the corresponding equation for the entropy tendency becomes

$$\dot{S} = P - Ar + \frac{\langle\eta^2\rangle}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial^2\rho}{\partial\tilde{q}^2}\ln\rho, \tag{D.16}$$

where we have reintroduced the non-local and non-linear terms via the advective production of entropy, P.

Thus, any increase in entropy must arise through the effect of turbulence or stochastic forcing. In the absence of turbulence, the stochastic noise term will solely act to spread out the probability distribution and will consequently act as a source of entropy. In this case we would expect there to be a stationary solution for the entropy evolution in the case of no eddies where the source of entropy due to stochastic noise balances the sink of entropy due to linear drag. This can be tested by numerically solving the linear equation of motion, (D.1), and comparing with the steady state solution of the Fokker-Planck equation, (D.2), and its corresponding entropy.

Appendix E

Time-evolution of maximal entropy

Suppose that entropy is maximized, constrained by time-varying energy and N polynomial Casimirs. The variational problem to be solved is given by

$$\begin{aligned} \frac{\delta S}{\delta \rho} &+ \alpha(t)\frac{\delta}{\delta \rho}\left(-\frac{1}{2}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho - E(t)\right) \\ &- \sum_{n=1}^{N}\gamma_n(t)\frac{\delta}{\delta \rho}\left(\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\tilde{q}^n\rho - C_n(t)\right) = 0. \end{aligned} \tag{E.1}$$

taking the functional derivative of the constraints we obtain

$$\frac{\delta S}{\delta \rho} = \alpha(t)\langle\psi\rangle\tilde{q} + \sum_{n=1}^{N}\gamma_n(t)\tilde{q}^n. \tag{E.2}$$

The factor of $\frac{1}{2}$ has disappeared from the first term because

$$\frac{\delta}{\delta \rho}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho = \frac{\delta}{\delta \rho}\iint \mathrm{d}^2\mathbf{x}\mathrm{d}^2\mathbf{x}'\mathrm{d}\tilde{q}\mathrm{d}\tilde{q}'\,\tilde{q}'\rho(\tilde{q}'|\mathbf{x}',t)G(\mathbf{x},\mathbf{x}')q\rho(q|\mathbf{x},t) \tag{E.3}$$

$$= 2\langle\psi\rangle\tilde{q} \tag{E.4}$$

where G is the Green's function for the potential vorticity inversion. To find the entropy tendency, we substitute this variational problem into the relation

$$\frac{dS}{dt} = \int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\frac{\delta S}{\delta\rho}, \tag{E.5}$$

giving

$$\frac{dS}{dt} = \int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\frac{\partial\rho}{\partial t}\left[\alpha(t)\langle\psi\rangle\tilde{q} + \sum_{n=1}^{N}\gamma_n(t)\tilde{q}^n\right]. \tag{E.6}$$

Rearranging and pulling out time derivatives we obtain

$$\frac{dS}{dt} = \frac{\alpha(t)}{2}\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho + \sum_{n=1}^{N}\gamma_n(t)\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\tilde{q}^n\rho, \tag{E.7}$$

and identifying the integrals with the constraints which they are, by construction, equal to, we obtain

$$\frac{dS}{dt} = -\alpha(t)\frac{dE(t)}{dt} + \sum_{n=1}^{N}\gamma_n(t)\frac{dC_n(t)}{dt}. \tag{E.8}$$

We note that because of the implicit dependence on time of $\langle\psi\rangle$ we have used the fact that

$$\frac{1}{2}\frac{d}{dt}\int \mathrm{d}^2\mathbf{x}\mathrm{d}\tilde{q}\,\langle\psi\rangle\tilde{q}\rho = \iint \mathrm{d}^2\mathbf{x}\mathrm{d}^2\mathbf{x}'\mathrm{d}\tilde{q}\mathrm{d}\tilde{q}'\,\tilde{q}'\rho(\tilde{q}'|\mathbf{x}',t)G(\mathbf{x},\mathbf{x}')q\frac{\partial\rho(q|\mathbf{x},t)}{\partial t}. \tag{E.9}$$

For ease of presentation we define $\alpha^* \equiv -\alpha$, completing the derivation for the entropy tendency

$$\frac{dS}{dt} = \alpha^*(t)\frac{dE(t)}{dt} + \sum_{n=1}^{N}\gamma_n(t)\frac{dC_n(t)}{dt}. \tag{E.10}$$

Appendix F

Significance testing

To exclude the possibility of statistical over-fitting of the entropy steady state fluctuations, in Section 6.5.2, we perform a significance test for the correlations observed. Firstly, cross-validate the regression by using half the data to train the regression while validating on the other half. We also create an ensemble of 'synthetic' time series which have the same power spectra as the diagnosed energy and Casimir time series. This is done using a MATLAB function, ebisuzaki.m, which is available as part of the WEACLIM toolbox by Vincent Moron, available on the MathWorks file exchange. This method follows the significance test of Ebisuzaki (1997). We test the case using energy and 10 Casimirs to reconstruct to entropy evolutions, we cross-validate both the reconstruction by diagnosed time series as well as the ensemble of reconstructions from 'synthetic' time series. The results are given in Figure F.1, where we have plotted the mean cross-validated correlation for the diagnosed time series, $\overline{c}$, given as

$$\overline{c} = \frac{c_1 + c_2}{2}, \tag{F.1}$$

where c_1 is the correlation between the reconstructed and diagnosed entropy for the second half of the data, trained on the first half; c_2 is defined vice versa. Also plotted are the percentiles for the ensemble of 1000 mean cross-validated correlations

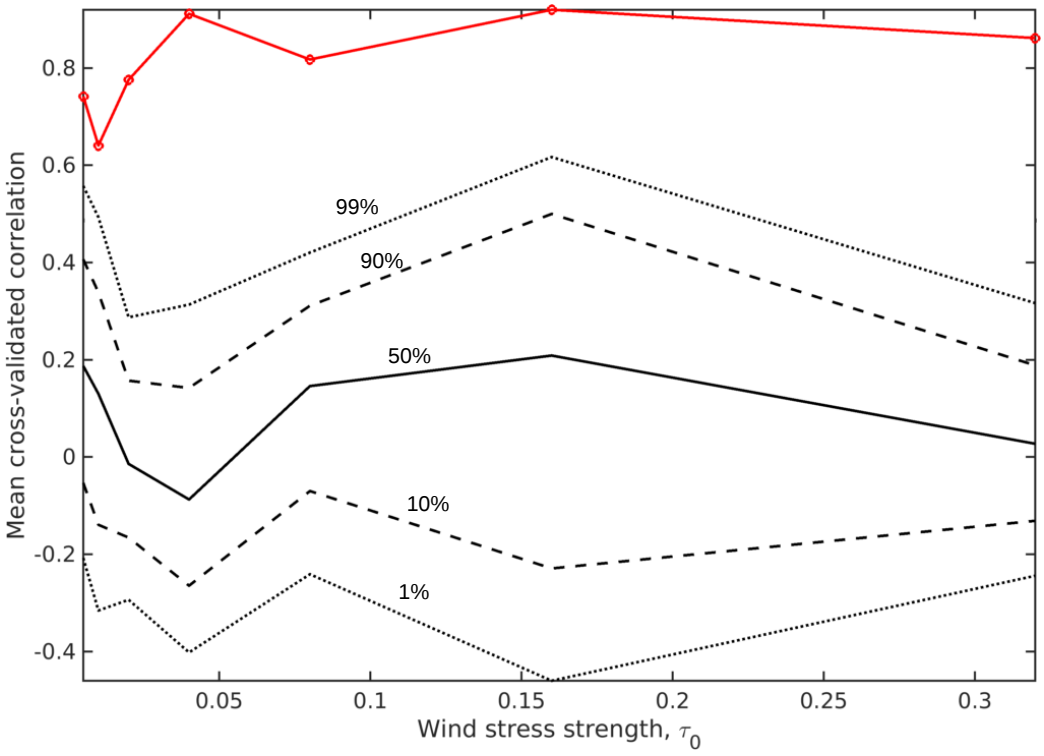


Figure F.1: *Red line* - mean cross-validated correlation between diagnosed and reconstructed entropy as a function of wind stress strength. *Black lines* - percentiles for the ensemble of mean cross-validated correlations from 'synthetic' time series of energy and Casimirs.

produced from the 'synthetic' time series. This clearly shows that the observed correlations lie beyond the 99th percentile of the ensemble and excludes the possibility of statistical over-fitting.

Appendix G

Reconstructing maximum entropy statistics

We have seen, from Chapter 6, equation (6.21), that the partition function can be written as

$$\mathcal{Z}(\mathbf{x}) = \int \mathrm{d}\tilde{q} \left[\frac{\Pi(\tilde{q}) e^{-\alpha\langle\psi\rangle\tilde{q}}}{\int \mathrm{d}^2\mathbf{x} \left(\frac{e^{-\alpha\langle\psi\rangle\tilde{q}}}{\mathcal{Z}(\mathbf{x})} \right)} \right], \tag{G.1}$$

which is an implicit equation for $\mathcal{Z}$. Writing $R(\tilde{q}, \mathbf{x}) = \exp(-\alpha\langle\psi\rangle\tilde{q})$ we can recast equation (G.1) as an iterative relation for the $i+1$ approximation for the partition function

$$\mathcal{Z}_{i+1}(\mathbf{x}) = \int \mathrm{d}\tilde{q} \left[\frac{\Pi(\tilde{q}) R(\tilde{q}, \mathbf{x})}{\int \mathrm{d}^2\mathbf{x} \left[\frac{R(\tilde{q}, \mathbf{x})}{\mathcal{Z}_i(\mathbf{x})} \right]} \right]. \tag{G.2}$$

Starting with $\mathcal{Z}_1 = 1$, we can produce an approximation for the partition function which normalizes the probability distribution, ρ, to a better than 1% accuracy after 20 iterations.

Using this methodology we are able to reduce the problem of determining the Lagrange multipliers to a one-dimensional problem of determining α by optimizing for the mean-field energy, E_M. Using a MATLAB numerical minimization function (`fminsearch`), we can optimize α in order to match the energy of the flow, that is,

by minimizing the cost function

$$L = \left| E_M - \frac{1}{2} \int \mathrm{d}^2\mathbf{x} \, \langle\psi\rangle \langle\tilde{q}\rangle \right|, \tag{G.3}$$

for given E_M, N and $\langle\psi\rangle$ which are diagnosed from the simulations.

Bibliography

Abernathey, R. and Marshall, J. (2013). Global surface eddy diffusivities derived from satellite altimetry. *Journal of Geophysical Research: Oceans*, 118:901–916.

Abernathey, R., Marshall, J., Mazloff, M., and Shuckburgh, E. (2010). Enhancement of mesoscale eddy stirring at steering levels in the southern ocean. *Journal of Physical Oceanography*, 40:170–184.

Abramov, R. V. and Majda, A. J. (2003). Statistically relevant conserved quantities for truncated quasigeostrophic flow. *PNAS*, 100(7):3841–3846.

Bartolotta, A., Carroll, S. M., Leichenauer, S., and Pollack, J. (2015). The Bayesian second law of thermodynamics.

Batchelor, G. K. (1969). Computation of the energy spectrum in homogeneous two-dimensional turbulence. *Physics of Fluids*, 12(II):233–239.

Beirlant, J., Dudewicz, E. J., Györfi, L., and van der Meulen, E. C. (1997). Nonparametric entropy estimation: an overview. *International Journal of Mathematical and Statistical Sciences*, 6(1):17–39.

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

Berrisford, P., Marshall, J., and White, A. (1993). Quasigeostrophic potential vorticity in isentropic coordinates. *Journal of the Atmospheric Sciences*, 50(5):778–782.

Bolton, S., Adriani, A., Adumitroaie, V., Allison, M., Anderson, J., Atreya, S., Bloxham, J., Brown, S., Connerney, J., DeJong, E., et al. (2017). Jupiters interior and deep atmosphere: The initial pole-to-pole passes with the juno spacecraft. *Science*, 356(6340):821–825.

Boucher, C., Ellis, R. S., and Turkington, B. (1999). Spatializing random measures: doubly indexed processes and the large deviation principle. *Annals of probability*, pages 297–324.

Bouchet, F. and Sommeria, J. (2002). Emergence of intense jets and jupiter's great red spot as maximum-entropy structures. *Journal of Fluid Mechanics*, 464:165–207.

Bouchet, F. and Venaille, A. (2012). Statistical mechanics of two-dimentional and geophysical flows. *Physics Reports*, 515:227–295.

Bourlioux, A. and Majda, A. J. (2002). Elementary models with probability distribution function intermittency for passive scalars with a mean gradient. *Physics of Fluids*, 14(2):881–897.

Bower, A. S., Rossby, H. T., and Lillibridge, J. L. (1985). The gulf stream - barrier or blender. *Journal of Physical Oceanography*, 15:24–32.

Bowman, J. C. (2013). Casimir cascades in two-dimensional turbulence. *Journal of Fluid Mechanics*, 729:364–376.

Bretherton, F. P. and Haidvogel, D. B. (1976). Two-dimensional trubulence above topography. *Journal of Fluid Mechanics*, 78:129–154.

Buizza, R., Miller, M., and Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. 125(560):2887–2908.

Carnevale, G. F. (1982). Statistical features of the evolution of two-dimensional turbulence. *Journal of Fluid Mechanics*, 122:143–151.

Carnevale, G. F., Frisch, U., and Salmon, R. (1981). H-theorems in statistical fluid dynamics. *J. Phys. A: Math. Gen.*, 14:1701–1718.

Chapman, C. C. (2014). Southern Ocean jets and how to find them: Improving and comparing common jet detection methods. *Journal of Geophysical Research: Oceans*, 119:4318–4339.

Charney, J. G. (1971). Geostrophic turbulence. *Journal of the Atmospheric Sciences*, 28:1087–1095.

Chavanis, P. H. (2009). Dynamical and thermodynamical stability of two-dimensional flows: variational principles and relaxation equations. *The European Physical Journal B*, 70(1):73–105.

David, T. W., Marshall, D. P., and Zanna, L. (2017). The statistical nature of turbulent barotropic jets. *Ocean Modelling*, 113:34–49.

Dijkstra, H. A. and Ghil, M. (2005). Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach. *Reviews of Geophysics*, 43(3).

Dritschel, D. G. and McIntyre, M. E. (2008). Multiple jets as PV staircases: The Phillips effect and the resilience of eddy-transport barriers. *J. Atmos. Sci.*, 65:855–874.

Dritschel, D. G., Qi, W., and Marston, J. B. (2015). On the late-time behaviour of a bounded, inviscid two-dimensional flow. *Journal of Fluid Mechanics*, 783:1–22.

Ebisuzaki, W. (1997). A method to estimate the statistical signifiance of a correlation when the data are serially correlated. *Journal of Climate*, 10:2147–2153.

Esler, J. G. (2008a). Robust and leaky transport barriers in unstable baroclinic flows. *Physics of Fluids*, 20(11):116602.

Esler, J. G. (2008b). The turbulent equilibration of an unstable baroclinic jet. *Journal of Fluid Mechanics*, 599:241–268.

Esler, J. G. and Haynes, P. H. (1999). Baroclinic wave breaking and the internal variability of the tropospheric circulation. *Journal of the Atmospheric Sciences*, 56(23).

Eyink, G. L. (1996). Exact results on stationary turbulence in 2D: consequences of vorticity. *Physica D*, 91:97–142.

Falkovich, G. and Hanany, A. (1993). Is 2D turbulence a conformal turbulence? *Phys. Rev. E*, 50:3454–3457.

Farrell, B. E. and Ioannou, P. J. (2003). Structural stability of turbulent jets. *Journal of the Atmospheric Sciences*, 60:2101–2118.

Farrell, B. F. and Ioannou, P. J. (2017). Statistical state dynamics based theory for the formation and equilibration of saturn's north polar jet. *Physical Review Fluids*, 2(7):073801.

Ferrari, R. and Nikurashin, M. (2010). Suppression of eddy diffusivity across jets in the southern ocean. *Journal of Physical Oceanography*, 40:1501–1519.

Fjørtoft, R. (1953). On the changes in the spectral distribution of kinetic energy for twodimensional, nondivergent flow. *Tellus*, 5(3):225–230.

Fox-Kemper, B., Bachman, S., Pearson, B., and Reckinger, S. (2014). Principles and advances in subgrid modelling for eddy-rich simulations. *Clivar Exchanges*, 65(2):42–46.

Frishman, A., Laurie, J., and Falkovich, G. (2017). Jets or vortices – What flows are generated by an inverse turbulent cascade? *Physical Review Fluids*, 2(3):032602.

Fultz, D., Long, R. R., Owens, G. V., Bohan, W., Keylor, R., and Weil, J. (1959). Studies of thermal convection in a rotating cylinder with some implications for large-scale atmospheric motions. In *Meteor. Monogr.*, volume 4, page 104. Amer. Meteor. Soc.

Gardiner, C. W. (2004). *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences*. Springer Series in Synergetics. Springer.

Garland, J., Jones, T. R., Bradley, E., James, R. G., and White, J. W. (2016). A first step toward quantifying the climates information production over the last 68,000 years. In *International Symposium on Intelligent Data Analysis*, pages 343–355. Springer.

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

Gent, P. R., Willebrand, J., McDougall, T. J., and Mcwilliams, J. C. (1995). Parameterizing eddy-induced tracer transports in ocean circulation models. *Journal of Physical Oceanography*, 25(4):463–474.

Gille, S. T., Speer, K., Ledwell, J. R., and Naveira Garabato, A. C. (2007). Mixing and stirring in the southern ocean. *Eos, Transactions American Geophysical Union*, 88(39):382–383.

Godfrey, D. (1988). A hexagonal feature around saturn's north pole. *Icarus*, 76(2):335–356.

Graham, R. M., de Boer, A. M., Heywood, K. J., Chapman, M., and Stevens, D. P. (2012). Southern ocean fronts: Controlled by wind or topography? *Journal of Geophysical Research*, 117.

Gritsun, A. and Branstator, G. (2007). Climate response using a three-dimensional operator based on the fluctuation–dissipation theorem. *Journal of the atmospheric sciences*, 64(7):2558–2575.

Grooms, I., Majda, A. J., and Smith, K. S. (2015). Stochastic superparameterization in a quasigeostrophic model of the Antarctic Circumpolar Current. *Ocean Modelling*, 85:1–15.

Hallberg, R. and Gnanadesikan, A. (2001). An exploration of the role of transient eddies in determining the transport of a zonally reentrant current. *Journal of Physical Oceanography*, 31(11):3312–3330.

Hallberg, R. and Gnanadesikan, A. (2006). The role of eddies in determining the structure and response of the wind-driven southern hemisphere overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) project. *Journal of Physical Oceanography*, 36:2232–2252.

Haller, G. and Beron-Vera, F. J. (2013). Coherent lagrangian vortices: The black holes of turbulence. *Journal of Fluid Mechanics*, 731.

Hannachi, A. (2007). Tropospheric planetary wave dynamics and mixture modelling: Two preferred regimes and a regime shift. *Journal of the Atmospheric Sciences*, 64(10):3521–3541.

Hasson, A., Koch-Larrouy, A., Morrow, R., Juza, M., and Penduff, T. (2012). The origin and fate of mode water in the southern pacific ocean. *Ocean Dynamics*, 62(3):335–354.

Hide, R. (1961). Origin of jupiter's great red spot. *Nature*, 190(4779):895.

Holm, D. D. (1986). Hamiltonian formulation of the baroclinic quasigeostrophic fluid equations. *Physics of Fluids (1958-1988)*, 29(1):7–8.

Holm, D. D. and Nadiga, B. T. (2003). Modeling mesoscale turbulence in the barotropic double-gyre circulation. *Journal of Physical Oceanography*, 33(11):2355–2365.

Holm, D. D. and Wingate, B. A. (2005). Baroclinic instabilities of the two-layer quasigeostrophic alpha model. *Journal of Physical Oceanography*, 35(7):1287–1296.

Hughes, C. W., Thompson, A. F., and Wilson, C. (2010). Identification of jets and mixing barriers from sea level and vorticity measurements using simple statistics. *Ocean Modelling*, 32:44–57.

Iselin, C. (1939). The influence of vertical and lateral turbulence on the characteristics of the waters at mid-depths. *Eos, Transactions American Geophysical Union*, 20(3):414–417.

Jackett, D. R. and McDougall, T. J. (1997). A neutral density variable for the world's oceans. *Journal of Physical Oceanography*, 27:237–263.

Jaynes, E. T. (1957a). Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630.

Jaynes, E. T. (1957b). Information theory and statistical mechanics. II. *The Physical Review*, 108(2):171–190.

Johnson, G. C. and Bryden, H. L. (1989). On the size of the Antarctic Circumpolar Current. *Deep-Sea Research*, 36(1):39–53.

Jung, S., Morrison, P. J., and Swinney, H. L. (2006). Statistical mechanics of two-dimensional turbulence. *Journal of Fluid Mechanics*, 554:433–456.

Karsten, R., Jones, H., and Marshall, J. G. (2002). The role of eddy transport in setting the stratification and transport of a circumpolar current. *Journal of Physical Oceanography*, 32:39–54.

Kaspi, Y., Galanti, E., Hubbard, W., Stevenson, D., Bolton, S., Iess, L., Guillot, T., Bloxham, J., Connerney, J., Cao, H., et al. (2018). Jupiters atmospheric jet streams extend thousands of kilometres deep. *Nature*, 555(7695):223.

Kazantsev, E., Sommeria, J., and Verron, J. (1998). Subgrid-scale eddy parameterization by statistical mechanics in a barotropic ocean model. *Journal of Physical Oceanography*, 28:1017–1042.

Klocker, A., Ferrari, R., and LaCase, J. H. (2012). Esimating suppression of eddy mixing by mean flows. *Journal of Physical Oceanography*, 42:1566–1576.

Kolmogorov, A. N. (1941). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Proceedings of the USSR Academy of Sciences*, 30:1417–1423.

Kraichnan, R. H. (1967). Inertial ranges in two-dimentional turbulence. *The Physics of Fluids*, 10(7):299–303.

Kraichnan, R. H. (1971). Inertial-range transfer in two- and three-dimensional turbulence. *Journal of Fluid Mechanics*, 47:525–535.

Kraichnan, R. H. and Montgomery, D. (1980). Two-dimentional turbulence. *Reports on Progress in Physics*, 43:547–619.

Kuo, H. I. (1949). Dynamic instability of two-dimensional non-divergent flow in a barotropic atmosphere. *J. Metreorol*, 6:105–122.

Landau and Lifshitz (1987). *Fluid Mechanics*, volume 6 of *Course of Theoretical Physics*. Butterworth Heinemann.

Lapeyre, G. and Klein, P. (2006). Impact of the small-scale elongated filaments on the oceanic vertical pump. *Journal of marine research*, 64(6):835–851.

Lapeyre, G., Klein, P., and Hua, B. L. (2006). Oceanic restratification forced by surface frontogenesis. *Journal of Physical Oceanography*, 36(8):1577–1590.

Le Traon, P., Nadal, F., and Ducet, N. (1998). An improved mapping method of multisatellite altimeter data. *Journal of atmospheric and oceanic technology*, 15(2):522–534.

Leith, C. (1975). Climate response and fluctuation dissipation. *Journal of the Atmospheric Sciences*, 32(10):2022–2026.

Leith, C. E. (1968). Diffusion approximation for two-dimensional turbulence. *Phys. Fluids*, 11:671–673.

Lévy, M. (2008). The modulation of biological production by oceanic mesoscale turbulence. In *Transport and Mixing in Geophysical Flows*, pages 219–261. Springer.

Lorenz, E. N. (1955). Available potential energy and the maintenance of the general circulation. *Tellus*, 7(2):157–167.

Lovenduski, N. S., Gruber, N., Doney, S. C., and Lima, I. D. (2007). Enhanced CO_2 outgassing in the Southern Ocean from a positive phase of the Southern Annular Mode. *Global Biogeochemical Cycles*, 21.

Lucarini, V., Fraedrich, K., and Lunkeit, F. (2010). Thermodynamic analysis of snowball earth hysteresis experiment: efficiency, entropy production and irreversibility. *Quarterly Journal of the Royal Meteorological Society*, 136(646):2–11.

Lucarini, V. and Sarno, S. (2010). A statistical mechanical approach for the computation of the climatic response to general forcings. *arXiv preprint arXiv:1008.0340*.

Lutsko, N. J., Held, I. M., and Zurita-Gotor, P. (2015). Applying the fluctuation–dissipation theorem to a two-layer model of quasigeostrophic turbulence. *Journal of the Atmospheric Sciences*, 72(8):3161–3177.

Lynden-Bell, D. (1967). Statistical mechanics of violent relaxation in stellar systems. *Mon. Not. R. Astr. Soc.*, 136:101–121.

Majda, A. and Wang, X. (2006). *Nonlinear dynamics and statistical theories for basic geophysical flows*. Cambridge University Press.

Majda, A. J. and Klein, R. (2003). Systematic multiscale models for the tropics. *Journal of the Atmospheric Sciences*, 60(2):393–408.

Marshall, D. P., Ambaum, M. H. P., Maddison, J. R., Munday, D. R., and Novak, L. (2017). Eddy saturation and frictional control of the antarctic circumpolar current. *Geophysical Research Letters*, 44(1):286–292. 2016GL071702.

Marshall, D. P. and Zanna, L. (2014). A conceptual model of ocean heat uptake under anthropogenic climate change. *Journal of Climate*.

Marshall, J., Jones, H., Karsten, R., and Wardle, R. (2002). Can eddies set ocean stratification? *Journal of Physical Oceanography*, 32:26–38.

Marshall, J., Shuckburgh, E., Jones, H., and Hill, C. (2006). Estimates and implications of surface eddy diffusivity in the southern ocean derived from tracer transport. *Journal of Physical Oceanography*, 36:1806–1821.

Marston, J. B., Conover, E., and Schneider, T. (2008). Statistics of an unstable barotropic jet from a cumulant expansion. *Journal of the Atmospheric Sciences*, 65:1955–1966.

Martyushev, L. M. and Seleznev, V. D. (2006). Maximum entropy production principle in physics, chemistry and biology. *Physics Reports*, 426:1–45.

Maze, G., Mercier, H., Fablet, R., Tandeo, P., Radcenco, M. L., Lenca, P., Feucher, C., and Le Goff, C. (2017). Coherent heat patterns revealed by unsupervised

classification of argo temperature profiles in the north atlantic ocean. *Progress in Oceanography*, 151:275–292.

Mazloff, M. R., Heimbach, P., and Wunsch, C. (2010). An eddy-permitting Southern Ocean state estimate. *Journal of Physical Oceanography*, 40:880–899.

McWilliams, J. C. (1977). A note on a consistent quasigeostrophic model in a multiply connected domain. *Dynamics of Atmospheres and Oceans*, 1:427–441.

Mead, L. R. and Papanicolaou, N. (1984). Maximum entropy in the problem of moments. *Journal of Mathematical Physics*, 25(8):2404–2417.

Merilees, P. E. and Warn, H. (1975). On energy and enstrophy exchanges in two-dimensional non-divergent flow. *Journal of Fluid Mechanics*, 69(4):625–630.

Merryfield, W. J. (1998). Effects of stratification on quasi-geostrophic inviscid equilibria. *Journal of Fluid Mechanics*, 354:345–356.

Miller, J. (1990). Statistical mechanics of euler equations in two dimensions. *Physical Review Letters*, 65(17):2137–2140.

Miller, J., Weichman, P. B., and Cross, M. C. (1992). Statistical mechanics, euler's equation, and jupiter's red spot. *Phys. Rev. A*, 45:2328–2359.

Montgomery, R. (1940). The present evidence on the importance of lateral mixing processes in the ocean. *Bulletin of the American Meteorological Society*, 21(3):87–94.

Morrison, P. J. (1998). Hamiltonian description of the ideal fluid. *Rev. Mod. Phys.*, 70:467–521.

Munday, D. R., Johnson, H. L., and Marshall, D. P. (2013). Eddy saturation of equilibrated circumpolar currents. *Journal of Physical Oceanography*, 43:507–532.

Munday, D. R., Johnson, H. L., and Marshall, D. P. (2014). Impacts and effects of mesoscale ocean eddies on ocean carbon storage and atmospheric pCO_2. *Global Biogeochemical Cycles*, 28:877–896.

Nakamura, N. (1996). Two-dimensional mixing, edge formation, and permeability diagnosed in an area coordinate. *J. Atmos. Sci.*, 53:1524–1537.

Nakamura, N. (2001). A new look at eddy diffusivity as a mixing diagnostic. *Journal of the atmospheric sciences*, 58(24):3685–3701.

Naso, A., Monchaux, R., Chavanis, P.-H., and Dubrulle, B. (2010). Statistical mechanics of Beltrami flows in axisymmetric geometry: Theory reexamined. *Physical Review E*, 81(6):066318.

Naviera Garabato, A. C., Ferrari, R., and Polzin, K. L. (2011). Eddy stirring in the Southern Ocean. *Journal of Geophysical Research*, 116.

Nazarenko, S. (2011). *Wave Turbulence*, volume 825 of *Lecture Notes in Physics*. Springer.

Nyquist, H. (1928). Thermal agitation of electric charge in conductors. *Physical review*, 32(1):110.

Onsager, L. (1949). Statistical hydrodynamics. *Nuovo Cimento*, 6(2):249–286.

Osborn, T. R. and Cox, C. S. (1972). Oceanic fine structure. *Geophysical and Astrophysical Fluid Dynamics*, 3(1):321–345.

Paltridge, G. W. (1975). Global dynamics and climate-a system of minimum entropy exchange. *Quarterly Journal of the Royal Meteorological Society*, 101(429):475–484.

Peixóto, J. P. and Oort, A. H. (1984). Physics of climate. *Reviews of Modern Physics*, 56(3):365.

Polyakov, A. M. (1993). The theory of turbulence in two dimensions. *Nuclear Physics B*, 396(2-3):367–385.

Porta Mana, P. G. L. and Zanna, L. (2014). Towards a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling*, 79.

Prieto, R. and Schubert, W. H. (2001). Analytical predictions for zonally symmetric equilibrium states of the stratospheric polar vortex. *Journal of the atmospheric sciences*, 58(18):2709–2728.

Qi, W. and Marston, J. B. (2014). Hyperviscosity and statistical equilibria of euler turbulence on the torus and the sphere. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(7):P07020.

Rayleigh, L. (1880). On the stability or instability of certain motions. *Proc. London Math. Soc.*, 11:57–70.

Renaud, A., Venaille, A., and Bouchet, F. (2016). Equilibrium statistical mechanics and energy partition for the shallow water model. *Journal of Statistical Physics*, 163(4):784–843.

Ring, M. J. and Plumb, R. A. (2008). The response of a simplified GCM to axisymmetric forcings: Applicability of the fluctuation–dissipation theorem. *Journal of the Atmospheric Sciences*, 65(12):3880–3898.

Rintoul, S., Hughes, C., and Olbers, D. (2001). *Ocean Circulation and Climate*, chapter The Antartic Circumpolar Current system. Academic Press.

Robert, R. (1990). Etats d'équilibre pour l'écoulement bidimensionnel d'un fluide parfait. *Comptes Rendus de l'Academie des Sciences 1*, 311:481–515.

Robert, R. (1991). A maximum-entropy principle for two-dimensional perfect fluid mechanics. *Journal of Statistical Physics*, 65:531–553.

Robert, R. and Sommeria, J. (1991). Statistical equilibrium states for two-dimensional flows. *Journal of Fluid Mechanics*, 229:291–310.

Ruelle, D. (1998). General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium. *Physics Letters A*, 245(3):220–224.

Ruelle, D. (2009). A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, 22(4):855.

Ryabko, B. and Reznikova, Z. (2009). The use of ideas of information theory for studying language and intelligence in ants. *Entropy*, 11(4):836–853.

Sallée, J. B., Speer, K., and Rintoul, S. R. (2011). Mean-flow and topographic control on surface eddy-mixing in the Southern Ocean. *Journal of Marine Research*, 69:753–777.

Salmon, R. (1998). *Lectures on Geophyical Fluid Dynamics*. Oxford University Press.

Salmon, R. (2012). Statistical mechanics and ocean circulation. *Commun. Nonlinear Sci. Numer. Simulat.*, 17:2144–2152.

Salmon, R., Holloway, G., and Hendershott, M. C. (1976). The equilibrium statistical mechanics of simple quasi-geostrophic models. *Journal of Fluid Mechanics*, 75:691–703.

Sasaki, H., Klein, P., Qiu, B., and Sasai, Y. (2014). Impact of oceanic-scale interactions on the seasonal modulation of ocean dynamics by the atmosphere. *Nature communications*, 5:6636.

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., and Williams, K. (2011). Improved Atlanic winter blocking in a climate model. *Geophysical Research Letters*, 38(L23703).

Schecter, D. A. (2003). Maximum entropy theory and the rapid relaxation of three-dimensional quasi-geostrophic turbulence. *Physical Review E*, 68(6):066309.

Schneider, T. and Griffies, S. M. (1999). A Conceptual Framework for Predictability Studies. *Journal of Climate*, 12.

Schneider, T., Teixeira, J., Bretherton, C. S., Brient, F., Pressel, K. G., Schär, C., and Siebesma, A. P. (2017). Climate goals and computing the future of clouds. *Nature Climate Change*, 7(1):3.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal*, 27(3):379–423.

Shuckburgh, E., Jones, H., Marshall, J., and Hill, C. (2009a). Robustness of an effective diffusivity diagnostic in oceanic flows. *Journal of Physical Oceanography*, 39(9):1993–2009.

Shuckburgh, E., Jones, H., Marshall, J., and Hill, C. (2009b). Understanding the regional variability of eddy diffusivity in the pacific sector of the southern ocean. *Journal of Physical Oceanography*, 39(9):2011–2023.

Sommeria, J. (2001). Two-dimensional turbulence. In *New trends in turbulence Turbulence: nouveaux aspects*, pages 385–447. Springer.

Sommeria, J., Meyers, S. D., and Swinney, H. L. (1989). Laboratory model of a planetary eastward jet. *Nature*, 337:58–61.

Stammer, D. and Wunsch, C. (1999). Temporal changes in eddy energy of the oceans. *Deep-Sea Res. II*, 46:77–108.

Stanley, G. J. (2017). The topology of neutral surfaces and their exact geostrophic stream function. *Manuscript in preparation*.

Starr, V. P. (1968). *Physics of negative viscosity phenomena*. McGraw-Hill.

Straub, D. N. (1993). On the transport and angular momentum balance of channel models of the antarctic circumpolar current. *Journal of Physical Oceanography*, 23(4):776–782.

Tansley, C. E. and Marshall, D. P. (2001). On the dynamics of wind-driven circumpolar currents. *Journal of Physical Oceanography*, 31:3258–3273.

Thalabard, S., Dubrulle, B., and Bouchet, F. (2014). Statistical mechanics of the 3D axisymmetric euler equations in a taylor–couette geometry. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(1):P01005.

Thalabard, S., Saint-Michel, B., Herbert, E., Daviaud, F., and Dubrulle, B. (2015). A statistical mechanics framework for the large-scale structure of turbulent von kármán flows. *New Journal of Physics*, 17(6):063006.

Thompson, A. F. and Demirov, E. (2006). Skewness of sea level variability of the world's oceans. *J. Geophys. Res.*, 111.

Thompson, A. F., Haynes, P. H., Wilson, C., and Richards, K. J. (2010). Rapid Southern Ocean front transitions in an eddy-resolving ocean GCM. *Geophys. Res. Lett.*, 37.

Treguier, A. M. (1989). Topographically generated steady currents in barotropic turbulence. *Geophysical and Astrophysical Fluid Dynamics*, 47:43–68.

Tung, K. K. and Welch, W. T. (2001). Remarks on charney's note on geostrophic turbulence. *Journal of the Atmospheric Sciences*, 58:2009–2012.

Turkington, B., Majda, A., Haven, K., and DiBattista, M. (2001). Statistical equilibrium predictions of jets and spots on jupiter. *Proceedings of the National Academy of Sciences*, 98(22):12346–12350.

Vallis, G. K. (2006). *Atmospheric and Oceanic Dynamics*. Cambridge University Press.

Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society, Series B*, 38:54–59.

Venaille, A. (2012). Bottom-trapped currents as statistical equilibrium states above topographic anomalies. *Journal of Fluid Mechanics*, 699:500–510.

Venaille, A. and Bouchet, F. (2011). Oceanic rings and jets as statistical equilibrium states. *Journal of Physical Oceanography*, 41:1860–1870.

Venaille, A., Gostiaux, L., and Sommeria, J. (2017). A statistical mechanics approach to mixing in stratified fluids. *Journal of Fluid Mechanics*, 810:554–583.

Venaille, A., Vallis, G. K., and Griffies, S. (2012). The catalytic role of the beta effect in barotropization processes. *Journal of Fluid Mechanics*, 709:490–515.

Volkov, D. L. and Zlotnicki, V. (2012). Performance of GOCE and GRACE-derived mean dynamic topographies in resolving Antarctic Circumpolar Current fronts. *Ocean Dynamics*, 62(6):893–905.

Wang, J. and Vallis, G. K. (1994). Emergence of Foffonoff states in inviscid and viscous ocean circulation models. *Journal of Marine Research*, 52:83–127.

Warn, T. (1986). Statistical mechanical equilibria of the shallow water equations. *Tellus A: Dynamic Meteorology and Oceanography*, 38(1):1–11.

Waterman, S., Hogg, N. G., and Jayne, S. R. (2011). Eddy-mean flow interactions in the kuroshio extension region. *Journal of Physical Oceanography*, 41:1182–1208.

Waterman, S. and Jayne, S. R. (2012). Eddy-driven recirculations from a localized transient forcing. *Journal of Physical Oceanography*, 42:430–447.

Waterman, S. and Lilly, J. M. (2015). Geometric decomosition of eddy feedbacks in barotropic systems. *Journal of Physical Oceanography*, 45:1009–1024.

Weichman, P. B. and Petrich, D. M. (2001). Statistical equilibrium solutions of the shallow water equations. *Physical Review Letters*, 86(9):1761.

Yasuda, Y., Bouchet, F., and Venaille, A. (2017). A new interpretation of vortex-split stratospheric sudden warmings in terms of equilibrium statistical mechanics. *arXiv preprint arXiv:1702.03716*.

Zakharov, V. and Piterbarg, L. (1988). Canonical variables for rossby waves and plasma drift waves. *Physics Letters A*, 126(89):497 – 500.

Zanna, L., Mana, P. P., Anstey, J., David, T., and Bolton, T. (2017). Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. *Ocean Modelling*, 111:66 – 80.

Zeitlin, V. (1991). On the structure of phase-space, hamiltonian variables and statistical approach to the description of two-dimensional hydrodynamics and magnetohydrodynamics. *Journal of Physics A: Mathematical and General*, 25:171–175.