Interactions between Near-Inertial Waves and Ocean Turbulence

Thesis by
Scott Conn

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2026
Defended Aug 6, 2025

Scott Conn
ORCID: 0000-0003-1404-0196

ACKNOWLEDGMENTS

"*E quindi uscimmo a riveder le stelle.*"

-Dante Alighieri

I thank Jörn Callies for sharing his encyclopedic knowledge of the ocean with me. He often knew where the project was headed long before I did, yet always left me room to puzzle things out on my own. His steady guidance and patience shaped my growth as an independent scientist, and he somehow taught me more about typography than I thought possible - after countless corrections of my stylistic missteps, I think I finally get it.

I am grateful to Andy Thompson for his encouragement and support over the last few years. Andy has a remarkable ability to see the ocean as an integrated system, and I hope some of that perspective has rubbed off on me. I always left our meetings with more motivation and confidence than I brought in.

I thank my committee members, Tapio Schneider and Leif Thomas, for their comments, advice, and most of all, their time. They pushed me to consider the broader implications of my work whenever I was too focused on the details. Tapio, in particular, deserves credit for urging me to incorporate the observational components of this thesis.

My co-authors, Joe Fitzgerald and Albion Lawrence, provided generous support throughout this project. When I began, I knew next to nothing about the ocean, and Joe's patience and guidance during those first steps proved indispensable. Working with Albion was so much fun, and reminded me of why I love theoretical physics. His ideas in this work were invaluable, but just as important were the wide-ranging chats that made working together such a pleasure.

At Caltech, I have been fortunate to work alongside many talented colleagues in Jörn and Andy's groups. The enthusiasm and generosity of current and former members made it an inspiring place to pursue a PhD. Their readiness to share ideas, discuss problems, and debug code taught me a great deal and spared countless hours of trial-and-error.

I could not have finished this thesis without the support of my friends. Luna, who sat by my side for the entire journey[1], was the best office-mate and friend I could have asked for; her steady encouragement kept me moving. Haakon, James, Calvin, Lacey, and Hemani were fantastic friends whose unwavering support – and all the adventures we squeezed in – made life outside the office something to look forward to. Will and Dan provided much of the entertainment that got me through grad school, and I'm grateful for the new experiences, their humour, and their relentless enthusiasm for life. There are too many others to list, both here in Pasadena and across Europe, to whom I am also extremely grateful.

[1] once COVID restrictions allowed us to sit next to people again

I have been fortunate to receive the support of my whole family, even across the miles that separate us. An extra-special thanks goes to Mum, Dad, and Ross – my biggest cheerleaders throughout this journey toward a PhD.

My greatest gratitude is reserved for Noah. Meeting him was, by far, the highlight of my PhD. His faith in my abilities often outstripped my own, and his countless pep talks and steady, practical help were indispensable in bringing this thesis to completion.

ABSTRACT

When atmospheric storms sweep across the sea surface they excite near-inertial waves (NIWs) whose frequency sits close to the local Coriolis frequency. By generating shear-driven mixing in the upper ocean, NIWs mediate the storm–ocean coupling and, in turn, influence climate. Oceanographers have long known that the fate of wind-generated NIWs hinges on how they interact with the background circulation – especially with mesoscale eddies – but observations show that this interaction ranges from negligible to dominant depending on location. Those regional contrasts matter: eddy-modified NIWs can dramatically reshape the pattern and intensity of mixing. My thesis uses a mix of theory, observations, and numerical modelling to better understand the dynamics that governs NIW-mesoscale interactions.

In trying to understand how NIWs behave in the presence of mesoscale eddies, the work of theoreticians culminated in the YBJ (Young and Ben-Jelloul) equation which describes the evolution of NIWs, including the physics of advection and refraction of NIWs by mesoscale eddies. I test whether this equation, subject to observations of the wind stress, stratification and mesoscale eddy field, can capture the observed dynamics of NIWs. Simulations of the YBJ equation can be compared to observations from a mooring array in the North Atlantic. The simulation reproduces the amplitude, phase, and across-array structure of the waves, and it reveals strong concentration of NIWs in anticyclones. In contrast, the traditional slab model – lacking mesoscale interaction physics – performs poorly. Potential energy budget diagnostics further show that, in this region, the net NIW–eddy energy exchange is small compared to other terms in the mesoscale energy budget.

Given the utility of the YBJ equation in understanding observations, I next try to understand theoretically what governs the impact that mesoscale eddies have on NIWs. This analysis heavily leverages the connection between the YBJ equation and the Schrödinger equation of quantum mechanics. The key governing parameter in the YBJ equation is the wave dispersiveness which quantifies the ratio of wave dispersion to wave refraction. Analytical calculations of the eigenmodes of the YBJ equation show that strongly dispersive waves are marginally affected by the eddies. However, eddies strongly imprint onto weakly dispersive NIWs. In the weak dispersion limit, the ray-tracing equations emerge from the YBJ equation, resolving some controversies regarding the applicability of ray-tracing to NIWs.

Finally, I try to understand how these different regimes may be distributed throughout the ocean. Observations from the Global Drifter Program can be used to calculate NIW spectra. Separating these spectra by vorticity reveals the impact of NIW-mesoscale interactions. NIW frequency shifts correlate strongly with vorticity, signalling weakly dispersive dynamics. Only a patch of the North Pacific shows a muted impact of mesoscale eddies. In high energy regions, such as western boundary currents and the Antarctic Circumpolar Current, NIWs exhibit a net negative frequency shift – a potential sign of strongly dispersive waves. The

true NIW signal is composed of many wave modes, each with a different dispersiveness, and each contributing to the observations. Idealised simulations of the YBJ equation are able to replicate the observed spectra from drifters well. The eigenmode approach is also useful in understanding the underlying physics that results in the observed spectral characteristics. The drifters confirm that trapping in anticyclones is common, but not universal.

PUBLISHED CONTENT AND CONTRIBUTIONS

Conn, S., Fitzgerald, J., and Callies, J., 2024: Interpreting Observed Interactions between Near-Inertial Waves and Mesoscale Eddies. *Journal of Physical Oceanography*, **54 (2)**, 485-502, doi:10.1175/JPO-D-23-0139.1.

All authors conceived the study. S.C. performed the data analysis and simulations. S.C. wrote the paper, and all authors contributed to the editing.

Conn, S., Callies, J., and Lawrence, A., 2025: Regimes of Near-Inertial Wave Dynamics. *Journal of Fluid Mechanics*, **1002**, A22, doi:10.1017/jfm.2024.1175.

All authors conceived the study. S.C. performed most of the calculations, J.C. performed the calculations of the invariant tori. S.C. wrote the paper, and all authors contributed to the editing.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

"*Physics has found no straight lines. Instead, the physical universe consists of only waves undulating back and forth allowing for corrections and balance.*"

-R. Buckminster Fuller

Humans have long been aware of an inextricable link between the winds and the waves. Over 2000 years ago, in his famous treatise on meteorology, *Meteorologica*, Aristotle hypothesised that the "winds are the causes of waves."[1] Although science has progressed enormously since the *Meteorologica*, his depiction of an ocean driven by the winds remains fundamentally correct. Some of Earth's strongest winds can be found over the ocean, where, unbridled by the constraints of surface friction, they can reach extraordinary speeds. On land we associate such winds with destruction, but over the sea they play a different role, engendering an energetic ocean circulation. Indeed, the winds are the ultimate source of the vast majority of kinetic energy in the ocean [2]. The global wind work has been estimated at 64 TW (Huang, 2004), with variability across a wide range of spatial and temporal scales. While most of this power ultimately fuels surface waves/turbulence, $\sim$4 TW gives rise to internal waves and the general circulation (Flexas et al., 2019).

Much as a trained opera singer can shatter glass by matching its natural frequency, variability in the winds excites vigorous internal motions when it resonates with the ocean's natural (Coriolis) frequency f. A power spectrum of the winds shows nothing remarkable at this frequency, yet a spectrum of the ocean velocity shows a peak at f, towering above the background of other waves (Figure 1.1). This peak is the hallmark of near-inertial waves (NIWs). Ubiquitous throughout the ocean, NIWs influence climate on both regional and global scales. This thesis aims to understand how NIWs interact with the rest of the turbulent ocean, motivated by a desire to understand the life-cycle of NIWs – from generation to dissipation – and the effects they have along the way.

NIWs are internal waves with a frequency close to f and represent a significant fraction of the high-frequency variability of the ocean. The strongest NIWs can reach up to $\sim$1 m s^{-1}. NIWs are circularly polarised, which arises from a balance between inertial and Coriolis terms in the momentum equation. NIWs are also associated with strong vertical shear, more so than other internal waves (see e.g., Pinkel, 2008). By virtue of being wind-generated,

[1] Aristotle almost certainly had no concept of internal waves, which will be the topic of this thesis. We'll give him a pass.

[2] A small but non-negligible portion of the ocean's kinetic energy is derived from the lunisolar tides (Wunsch and Ferrari, 2004).

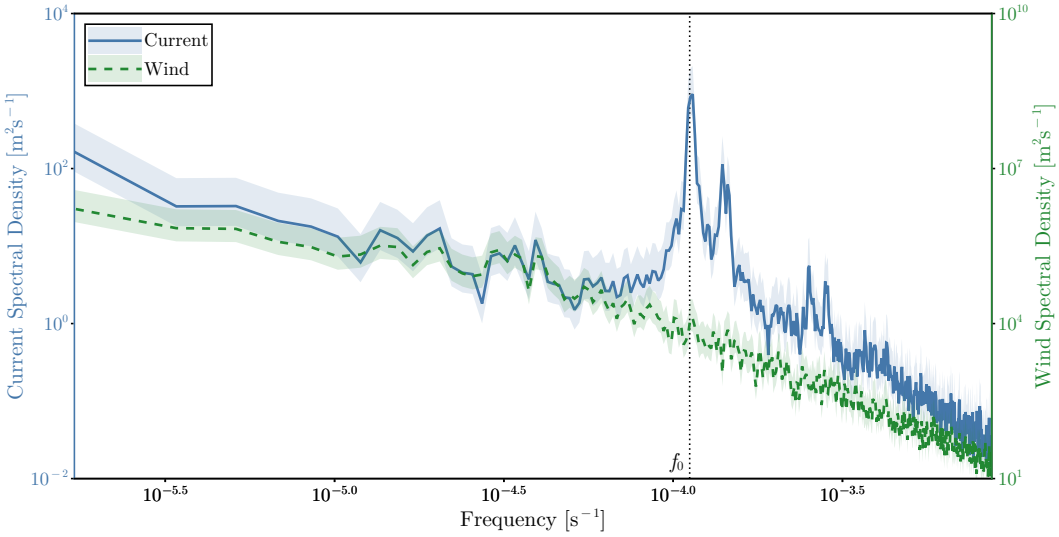

Figure 1.1: Figure showing the power spectrum of the ocean current velocity (solid blue) compared to the power spectrum of the wind velocity (dashed green). The black dotted line shows the location of the Coriolis frequency f. Data is taken from the mooring at Ocean Station Papa in the North Pacific (50.1°N). The data is provided at hourly resolution from 2007 to 2025, and split into 42 day segments to calculate the spectra.

NIWs show up as an intermittent signal in observations. These unique features of NIWs lead to important impacts on the climate but also make them hard to observe.

Previous work has highlighted a myriad of important roles that NIWs play in the global climate system. The strong vertical shear of NIWs can generate shear instabilities. As a result, NIWs deepen the mixed-layer, especially in the wake of atmospheric storms (Plueddemann and Farrar, 2006; Alford, 2020). Improperly representing NIWs in a model can result in biases in sea-surface temperature, precipitation and atmospheric winds due to biases in the mixed-layer depth (Jochum et al., 2013). In the Arctic, NIWs have been implicated in the mixing of heat from the deeper, warmer Atlantic water into the surface (Fer, 2014). Near 30°, the diurnal frequency corresponds with the local Coriolis frequency and strong NIWs can be generated by the diurnal breeze. These strong NIWs and their interactions with currents have been shown to influence the ventilation of deeper waters (Qu et al., 2022) and phytoplankton dynamics (Lucas et al., 2014). NIWs are also an important feature of the ocean's response to tropical cyclones. The strong winds associated with tropical cyclones can excite strong NIWs which can then contribute to enhanced vertical mixing (see e.g., Gutiérrez Brizuela et al., 2023; Yuan et al., 2024; Lazaneo et al., 2024).

Unfortunately, NIWs are hard to observe and we are consequently lacking a clear global picture of their dynamics. NIWs have no leading-order signal in sea-surface height (recall that the primary balance for NIWs does not involve the pressure gradient term) and so NIWs are invisible to satellite altimetry. Therefore, it remains unclear where or how much energy is entering and leaving the NIW band in a global sense. We are left with local *in-situ* measure-

ments and modelling to understand the dynamics of NIWs. Therefore, to fully understand the impacts of NIWs on the global climate we need to fully leverage available observations and theory to help us understand the propagation and dynamics of NIWs.

Wind-generated NIWs do not propagate through a quiescent ocean; instead they are forced into an already turbulent ocean. This flow consists of large-scale, mesoscale, and submesoscale motions as well as tides and other internal waves. In the past few decades it has become clear that interactions with the mesoscale can strongly govern the evolution of NIWs. Observational studies reveal that the importance of the mesoscale to NIW dynamics likely varies spatially in the ocean. Some observational studies show NIWs whose evolution is strongly constrained by the presence of mesoscale eddies (see e.g., Thomas et al., 2020). These studies highlight the rich dynamics of NIW-mesoscale interactions. A symptom of this interaction is that NIWs may become concentrated into regions of anticyclonic vorticity. This provides an efficient route for NIWs to propagate out of the mixed-layer [3] in a phenomenon that has been termed the inertial chimney (Lee and Niiler, 1998) or drainpipe (Asselin and Young, 2020). Other observational studies show a more muted role of NIW-mesoscale interactions in the evolution of NIWs (see e.g., D'Asaro et al., 1995). While these observational studies are crucial to understanding the dynamics of NIWs, their interpretation has been greatly aided by progress in our theoretical understanding of NIW-mesoscale interactions.

Theoretical progress relies on making approximations to the equations of motion. Early studies of NIWs were based on simple slab mixed-layer models (Pollard and Millard Jr, 1970). These models retain the inertial, Coriolis and wind-stress terms in the horizontal momentum equations. All of the physics that results in propagation of NIWs out of the mixed-layer is then parametrised in the model as a linear decay. This is a drastic simplification of NIW dynamics, yet the slab mixed-layer model remains a popular tool to simulate NIWs given the practicality of running such models across the entire globe (Alford, 2001). Kunze (1985) developed a theory that tried to better resolve the physics of the interaction. By assuming the waves are propagating through a slowly-varying background flow, ray-tracing can be used to calculate the path of NIW packets through a balanced background field. To perform ray-tracing calculations, Kunze (1985) derived a dispersion relation for NIWs in the presence of a background field [4]:

$$\omega = f + \frac{1}{2}\zeta + \frac{N^2 K^2}{2 f_0 m^2} + \boldsymbol{k} \cdot \boldsymbol{U}, \tag{1.1}$$

where ω is the wave frequency, $\boldsymbol{U} = (U, V)^T$ is the background velocity, $\zeta = \partial_x V - \partial_y U$ is the background vorticity, N^2 is the stratification, $\boldsymbol{k} = (k, l, m)^T$ is the wavevector, and $K = \sqrt{k^2 + l^2}$ is the horizontal wavenumber. The ray-tracing approach has come under criticism for its spatial-scale separation assumption. In particular, if NIWs are forced by large-scale

[3] The vertical group velocity of NIWs scales as the horizontal wavenumber squared. Concentration into anticyclones reduces the lateral scale of the NIWs and hence increases the vertical group velocity.

[4] The equation shown here is simplified by assuming a barotropic mean flow. The full dispersion relation can be found in Kunze (1985) - Equation 6.

atmospheric storms, they would be propagating through a quickly varying medium. Young and Ben-Jelloul (1997, from hereon YBJ) developed a theory to explain NIW propagation through a mesoscale eddy field that relies only on the temporal scale separation between the fast waves and the slowly evolving mesoscale flow. The YBJ equation is an evolution equation for the NIW field [5]:

$$\frac{\partial^3 M}{\partial^2 z \partial t} + J\left(\psi, \frac{\partial^2 M}{\partial^2 z}\right) + \frac{iN^2}{2f}\nabla^2 M + \frac{i\zeta}{2}\frac{\partial^2 M}{\partial^2 z} = 0, \tag{1.2}$$

where $\boldsymbol{u} = (u, v)^T$ is the NIW horizontal velocity, $\partial_z M = (u + iv)e^{if_0 t}$ is the back-rotated and complexified NIW velocity, $J(a, b) = \partial_x a \cdot \partial_y b - \partial_x b \cdot \partial_y a$ is the Jacobian operator, and ψ is the mesoscale streamfunction. In this equation, the NIW-mesoscale interaction shows up in the advection term as well as in the term involving $\zeta = \nabla^2\psi$. Xie and Vanneste (2015, from hereon XV) extended the YBJ theory to include the back-reaction of the NIWs on the mesoscale eddy field. The XV model predicts that NIW potential energy should be generated at the expense of mesoscale kinetic energy in a process known as stimulated generation. The YBJ/XV models represent the most complete theoretical models of NIW-mesoscale interactions that we have. However, they still require input from observations/high-resolution numerical models to quantify the relative importance of the various physical processes that they predict. In Chapter 2 of my thesis I use these models to better understand the processes that govern NIW evolution in the real ocean and how they may manifest themselves in observations.

The YBJ equation 1.2 is a complicated 3D partial differential equation. Given that we are assuming a barotropic background flow, we can obtain a simplification by decomposing the solution into the vertical baroclinic modes. Writing the NIW velocity for a single baroclinic mode as $[u_w(x, y, t), v_w(x, y, t)]g(z)$, with $g(z)$ being the baroclinic mode structure, we can define a back-rotated velocity as $\phi = (u_w + iv_w)e^{if t}$. The YBJ equation then becomes

$$\frac{\partial \phi}{\partial t} + J(\psi, \phi) + \frac{i\zeta}{2}\phi - \frac{if\lambda^2}{2}\nabla^2\phi = 0, \tag{1.3}$$

where λ is the deformation radius associated with the baroclinic mode. There is one non-dimensional number associated with this equation known as the wave dispersiveness. If Ψ is the scale of the streamfunction then the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$. It represents the ratio of the strength of dispersion to the strength of refraction. Thomas et al. (2024a) argued that differences in the wave dispersiveness were partially responsible for the different importance of mesoscale eddies to the evolution of NIWs in different regions of the ocean. In Chapter 3 of my thesis, I apply techniques from quantum mechanics to this form of the YBJ equation to understand how NIWs behave in these different regimes. Using this framework, I will also show how the ray-tracing framework of Kunze (1985) relates to the YBJ equation.

[5] We again show the equation for a barotropic mean flow. We also set the planetary vorticity gradient β to zero. The full equation can be found in Young and Ben-Jelloul (1997) - Equation 4.7.

Compared to these theoretical models, NIWs in the real ocean display a great deal of complexity. Differences in the properties of the atmospheric winds and the structure of the mixed layer influence the amount and vertical structure of near-inertial energy forced by the winds. Differences in the local stratification and mesoscale eddy field will influence the subsequent propagation and dissipation of NIWs. The variability of NIW dynamics, and in particular the variability in the interaction between NIWs and mesoscale eddies, has not been well-mapped at a global scale. Such a mapping is essential to understand the role of NIWs in the climate system, as many of their impacts depend sensitively on the details of how the NIWs evolve, which in turn depends on the characteristics of the NIW-mesoscale interaction. Using data from the Global Drifter Program, Elipot et al. (2010) found evidence that NIW spectral properties are modulated by mesoscale eddies globally. When averaging over all drifter observations, they showed that the frequency of the NIW peak is shifted by $\sim 0.4\zeta$ – close to the ray-tracing prediction of $\zeta/2$. However, such a global mean cannot provide any information about how the characteristics of the NIW-mesoscale interactions are distributed throughout the ocean. In Chapter 4 of my thesis, I diagnose NIW spectral characteristics [6] and their dependence on the mesoscale eddy field with global resolution. I show how these characteristics are expected from the YBJ model and explain the underlying physics that gives rise to them.

[6] The stated motivation focuses on understanding the role of NIWs in the climate system. However, this is also a pure physics question. If we claim to understand the dynamics of NIWs, then we should be able to explain what sets the difference in their spectral properties throughout the ocean.

Chapter 2

INTERPRETING OBSERVED INTERACTIONS BETWEEN NEAR-INERTIAL WAVES AND MESOSCALE EDDIES

This chapter is reproduced from the published article:

Conn, S., Fitzgerald, J., and Callies, J., 2024: Interpreting Observed Interactions between Near-Inertial Waves and Mesoscale Eddies. *Journal of Physical Oceanography*, **54 (2)**, 485-502, doi:10.1175/JPO-D-23-0139.1.

As a self-contained work, some notation may differ from conventions used elsewhere in this thesis.

2.1 Abstract

The evolution of wind-generated near-inertial waves (NIWs) is known to be influenced by the mesoscale eddy field, yet it remains a challenge to disentangle the effects of this interaction in observations. Here, the model of Young and Ben Jelloul (YBJ), which describes NIW evolution in the presence of slowly evolving mesoscale eddies, is compared to observations from a mooring array in the Northeast Atlantic Ocean. The model captures the evolution of both the observed NIW amplitude and phase much more accurately than a slab mixed layer model. The YBJ model allows for the identification of specific physical processes that drive the observed evolution. It reveals that differences in the NIW amplitude across the mooring array are caused by the refractive concentration of NIWs into anticyclones. Advection and wave dispersion also make important contributions to the observed wave evolution. Stimulated generation, a process by which mesoscale kinetic energy acts as a source of NIW potential energy, is estimated to be 20 μW m^{-2} in the region of the mooring array, which is two orders of magnitude smaller than the global average input to mesoscale kinetic energy and likely not an important contribution to the mesoscale kinetic energy budget in this region. Overall, the results show that the YBJ model is a quantitatively useful tool to interpret observations of NIWs.

2.2 Introduction

Near-inertial waves (NIWs), internal waves with a frequency close to the inertial frequency f, are resonantly excited by atmospheric winds exerting a stress on the ocean's surface. It has long been recognized that these waves can interact with mesoscale eddies and that this interaction may be important in the life cycle of wind-generated NIWs. Observational evidence of NIW–mesoscale interactions is accumulating, but interpreting the observed NIW evolution in the presence of mesoscale eddies remains challenging. Here, we employ the theoretical

framework of Young and Ben-Jelloul (1997, from hereon YBJ) to identify and interpret such interactions in mooring observations in the Northeast Atlantic Ocean.

NIWs are associated with a large vertical shear, which can result in shear instabilities and vertical mixing (for a review, see Alford et al., 2016). These shear instabilities are a key mechanism by which atmospheric storms can cause the surface mixed layer to deepen. Jochum et al. (2013) showed that an improved representation of NIWs in a climate model led to a deepening of the mixed layer on average, which in turn resulted in significant changes in sea surface temperatures, winds, and precipitation.

The amount of NIW shear in the upper ocean, and as a consequence the amount of mixed layer deepening, depends on both the energy input into the NIW band and how rapidly NIWs propagate to depth. The vertical propagation originally presented a paradox. The vertical group velocity of NIWs varies as κ^2, where κ is the horizontal wavenumber. The atmospheric storms that generate NIWs are typically $O(1000\ \mathrm{km})$ in size and generate NIWs with a similar scale. Estimates of the vertical group velocity based on this horizontal scale are much too slow to explain the observed decay of NIWs in the mixed layer and propagation to depth (D'Asaro et al., 1995). A reduction in the horizontal scale of the waves is required to obtain a group velocity that matches observations. There are two possible mechanisms by which this is thought to occur: (i) meridional variations in f cause an increase in the meridional wavenumber (β-refraction; Gill, 1984), or (ii) interactions with the mesoscale circulation can imprint structure onto the wave field at the scale of mesoscale eddies (ζ-refraction; YBJ; Kunze, 1985). The presence of the former alone is enough to cause propagation of NIWs out of the mixed layer, with the NIW kinetic energy in the mixed layer decaying as $t^{-3/2}$ (Moehlis and Smith, 2001). The latter process causes concentration of NIW energy into anticyclones and subsequent downward propagation in inertial "chimneys" (Lee and Niiler, 1998) or "drainpipes" (Asselin and Young, 2020).

There is a long history of NIW observations from *in situ* measurements. The Ocean Storms Experiment (D'Asaro et al., 1995) was a groundbreaking observational campaign to study the life cycle of wind-generated NIWs. The experiment tracked the response of the surface ocean to strong wind forcing in a region with weak mesoscale eddies. It was found that β-refraction was the dominant process driving the observed evolution of the meridional wavenumber of the waves. The process of ζ-refraction was found to have much less of an effect than expected (D'Asaro, 1995), which YBJ argued was due to strong dispersion in a weak eddy field. More recently, Thomas et al. (2020) used ship-based observations of NIWs in a strong dipole vortex to estimate the rate of change in time of the NIW horizontal wavenumbers. Here, the expected change in horizontal wavenumbers from ζ-refraction was consistent with the data. The observations further showed a NIW beam at depth, indicating that the scale reduction had allowed the NIWs to propagate vertically out of the mixed layer. Essink et al. (2022) studied typhoon-forced NIWs in the Kuroshio Current. They observed

the trapping of NIWs in a region of anticyclonic vorticity, followed by downward propagation. They also measured the vertical structure of turbulence and showed how this was influenced by NIW dynamics. Yu et al. (2022a) investigated the interaction of NIWs and mesoscale eddies in observational data of NIWs from the Northeast Atlantic (the same data as we will use below). They showed that regions of elevated NIW kinetic energy are statistically associated with mesoscale anticyclones and that the submesoscale vorticity exerted little control on the horizontal concentration of NIWs and the subsequent propagation to depth. These studies, however, often focus on a single physical process (usually refraction) and many make assumptions about the NIW dynamics that may not be universally justified. Here, we propose that the YBJ model is a general theoretical framework that, when applied to observations, allows us to understand the multiple physical processes that govern NIW evolution.

It has also been proposed that NIWs not only react to the presence of mesoscale eddies but feed back on the eddies and affect their evolution. Approximately 80% of the ocean's kinetic energy exists as mesoscale motions (Ferrari and Wunsch, 2009). The geostrophic constraint on mesoscale eddies traps energy at large scales, and it is not entirely clear how the energy input into mesoscale motion is balanced by dissipation (Müller et al., 2011). A number of mechanisms by which mesoscale eddies lose energy are known, including dissipation in bottom boundary layers (e.g., Arbic and Flierl, 2004), the generation of dissipative lee waves (e.g., Nikurashin et al., 2013), energy loss near western boundaries (e.g., Zhai et al., 2010), and the top drag arising from the current dependence of the wind stress (e.g., Dewar and Flierl, 1987; Renault et al., 2016). The extraction of energy from mesoscale eddies by NIWs presents another possibility (Xie and Vanneste, 2015; Rocha et al., 2018). Given the great importance of mesoscale eddies to the transport of heat and carbon (e.g., Jayne and Marotzke, 2002; Gnanadesikan et al., 2015), even small changes (see discussion in Asselin and Young, 2020) to the mesoscale eddy field caused by NIWs may be significant to the impact of the ocean on climate.

To understand the propagation of NIWs and by extension the role they play in upper-ocean mixing, we need to understand the dynamics governing their evolution. Slab mixed layer models (Pollard and Millard Jr, 1970, from hereon PM) are a commonly used tool to model NIW evolution (e.g., D'Asaro, 1985; Alford, 2001, 2020; Guan et al., 2014). PM showed that a slab model can reproduce key features of the NIW evolution observed by moorings. With the horizontal NIW velocity denoted by $(u, v)^{\mathrm{T}}$, the PM model can be written as:

$$\partial_t u - fv = -ru + \frac{\tau^x}{\rho_w H_m}, \tag{2.1a}$$

$$\partial_t v + fu = -rv + \frac{\tau^y}{\rho_w H_m}, \tag{2.1b}$$

where r^{-1} is a decay timescale, ρ_w is the density of water, H_m is the mixed layer depth, and $\boldsymbol{\tau}$ is the wind stress. This model does not explicitly represent any physical processes affecting

the evolution of NIWs other than the wind forcing and Coriolis effect. All other processes are subsumed in the linear drag, which can be thought of as a parameterization primarily for the propagation of NIWs out of the mixed layer. Observations indicate that this process is slow relative to the inertial frequency (e.g., D'Asaro et al., 1995), and so we require $r \ll f$. Other than this restriction, r is a tunable parameter.

Despite its successes in capturing some aspects of the observed NIW signal, the PM model cannot explain the propagation of NIWs out of the mixed layer. If the NIW field is initially uniform, it will remain so. The model captures neither β-refraction nor the interaction with mesoscale eddies.

A more general framework to understand the evolution of NIWs was devised by YBJ by assuming a time scale separation between the fast waves and the slowly evolving mesoscale flow. In the YBJ framework, the horizontal NIW velocity is first complexified (written as $u + iv$). Since NIWs have a frequency close to f, it is convenient to write this complexified velocity as $u + iv = e^{-ift}\partial_z M(x, y, z, t)$. The function $\partial_z M$ describes the slow evolution of the envelope that modulates the NIW phase and amplitude. The YBJ equation describes how $\partial_z M$ evolves in the presence of prescribed geostrophic mesoscale eddies. On an f-plane, the equation reads

$$\partial_{zzt} M + J(\psi, \partial_{zz} M) + \frac{iN^2}{2f}\nabla^2 M + \frac{i\zeta}{2}\partial_{zz} M = \partial_{zz}\mathcal{F} - \nu\nabla^4\partial_{zz} M, \tag{2.2}$$

where ψ is the geostrophic streamfunction, $\nabla^2 = \partial_x^2 + \partial_y^2$ is the horizontal Laplacian operator, $\zeta = \nabla^2\psi$ is the geostrophic vorticity, N^2 is the stratification, $\mathcal{F}$ is a forcing term that represents the momentum flux due to the surface wind stress, and ν is a hyper-diffusivity included for numerical stability (see Section 2.4). In (2.2) and throughout this paper, we assume the geostrophic flow to be barotropic (i.e., ψ is independent of depth), although baroclinicity in the geostrophic eddy field can also be taken into account. Unlike the PM model, the YBJ equation does not have a tuneable parameter.

The second term in (2.2) represents advection of NIWs by the mesoscale flow. The third term represents changes in the NIW field due to dispersion. The presence of the dispersion term means that an initially localized wave packet will spread out as time progresses. The fourth term is responsible for the process of ζ-refraction. This term sets into motion the imprinting of mesoscale structure onto an initially horizontally uniform wave field. This ζ-refraction shifts the phase of the NIWs, which we can see by neglecting all other terms in the YBJ equation:

$$\partial_{zzt} M = -\frac{i\zeta}{2}\partial_{zz} M. \tag{2.3}$$

Assuming a steady vorticity field yields solutions of the form $\partial_{zz} M = C(x, y, z)e^{-i\zeta t/2}$. Spatial heterogeneities in the mesoscale vorticity will result in spatial heterogeneities in the NIW phase. The dispersion term in the full YBJ equation (2.2) then acts on these phase

gradients and fluxes energy into anticyclones and out of cyclones. This can be seen from the YBJ kinetic energy budget (Appendix A):

$$\partial_t \mathcal{K} + J(\psi, \mathcal{K}) + \nabla \cdot \mathbf{F}_{\mathcal{K}} + \partial_z G_{\mathcal{K}} = \gamma^F_{\mathcal{K}} + d_{\mathcal{K}}, \tag{2.4}$$

where $\mathcal{K}$ is the NIW kinetic energy density, $\mathbf{F}_{\mathcal{K}}$ and $G_{\mathcal{K}}$ are the horizontal and vertical energy fluxes due to dispersion, $\gamma^F_{\mathcal{K}}$ is the energy input by the forcing, and $d_{\mathcal{K}}$ is the dissipation due to the hyperviscosity term. The horizontal energy flux $\mathbf{F}_{\mathcal{K}}$ can be expressed in terms of gradients of the phase Θ of M (Rocha et al., 2018):

$$\mathbf{F}_{\mathcal{K}} = \frac{N^2 |M|^2}{2f} \nabla \Theta. \tag{2.5}$$

The spatial heterogeneities that the mesoscale vorticity ζ induces in the NIW phase thus cause a transfer of NIW energy in the horizontal. Once this ζ-refraction has imprinted the horizontal structure of eddies onto an initially uniform NIW field, advection will also act on the resulting gradients and stir the NIW field. Dispersion remains important and helps waves escape straining regions (Rocha et al., 2018). The scale reduction also accelerates the propagation of NIWs out of the mixed layer. This is represented in the YBJ energy budget (2.4) through the vertical flux $G_{\mathcal{K}}$, which requires horizontal gradients in the NIW field to produce energy fluxes to depth (Appendix A).

The energy input into the NIW band by the winds $\gamma^F_{\mathcal{K}}$ is known as the NIW wind work. The difference between the NIW wind work and the NIW kinetic energy that propagates out of the surface layer is the energy that is available for NIW mixing in the surface ocean. Both processes therefore influence the mixed layer depth. There has been extensive effort to estimate the NIW wind work. It can be calculated directly from concurrent observations of winds and NIW surface currents, but this is possible in a few locations only. Slab models have been used to obtain global estimates of the NIW wind work (Alford, 2001) but these are suspected to be overestimates, primarily because these models poorly represent the various processes that cause NIWs to leave the mixed layer. This can include the vertical propagation discussed above, but the employed models often also do not represent dissipation of NIWs by conversion to turbulent kinetic energy, which would reduce the projection of the wind stress onto the waves (Plueddemann and Farrar, 2006; Alford, 2020). High-resolution ocean models can also be used to estimate the wind work, but here problems can arise from limitations in the reanalysis data used to force the models. For example, Flexas et al. (2019) showed that the NIW wind work was poorly represented in a high-resolution ocean model due to the wind forcing missing variability on scales less than 6 hours and 15 km. The power input is larger in high-resolution coupled simulations but still substantially lower than observational estimates from slab models (von Storch and Lüschow, 2023).

The mesoscale can act as a source of wave potential energy in a process known as stimulated generation. In unforced and inviscid YBJ dynamics, the wave kinetic energy is conserved

but NIWs can gain or lose potential energy through interactions with the mesoscale. In the YBJ framework, the domain-integrated potential energy is not conserved (Appendix B):

$$\partial_t \langle \mathcal{P} \rangle = \Gamma^R_{\mathcal{P}} + \Gamma^A_{\mathcal{P}} + \mathcal{D}_{\mathcal{P}}, \tag{2.6}$$

where $\mathcal{P}$ is the potential energy density, $\Gamma^R_{\mathcal{P}}$ is the production of potential energy by refraction, $\Gamma^A_{\mathcal{P}}$ is the production of potential energy by advection, and $\mathcal{D}_{\mathcal{P}}$ is the potential energy dissipation by hyperviscosity. The formation of horizontal structure in an initially uniform NIW causes an increase in NIW potential energy via $\Gamma^R_{\mathcal{P}}$ and $\Gamma^A_{\mathcal{P}}$. In the first phase of interaction, $\Gamma^R_{\mathcal{P}}$ dominates as the waves have the mesoscale structure imprinted on them. At later times, $\Gamma^A_{\mathcal{P}}$ becomes the most important term. An extension of YBJ describing the coupled evolution of NIWs and a quasi-geostrophic mesoscale field was derived by Xie and Vanneste (2015) (see also Wagner and Young, 2016). This extension shows that the refractive and advective sources of potential energy to the NIWs appear as sinks in the mesoscale energy budget. This allows us to interpret $\Gamma^R_{\mathcal{P}}$ and $\Gamma^A_{\mathcal{P}}$ as energy transfers from the mesoscale, despite the fact that we impose the mesoscale field in the simulations described below. Theoretical and numerical studies (Xie and Vanneste, 2015; Rocha et al., 2018; Asselin and Young, 2020) have investigated the process of stimulated generation, yet its importance in the real ocean remains poorly constrained.

Several attempts have been made to reconcile available observations with our theoretical understanding of NIW–mesoscale interactions. Work on the NIW–mesoscale interaction prior to YBJ had been based on ray tracing theory, which additionally assumes that the waves have much shorter spatial scales than the background mesoscale flow (Kunze, 1985). Ray tracing predicts the NIW frequency to be shifted by $\frac{1}{2}\zeta$. This prediction, however, applies only in regions of the ocean where the WKB limit of ray tracing is appropriate. YBJ argued that this was not the case for the Ocean Storms Experiment, suggesting that the observations were taken in a region where the waves are instead in the so-called strong-dispersion limit. In this limit, refraction of the large-scale wave field is strongly opposed by dispersion, and the frequency shift is much smaller than predicted by WKB theory. This provides a compelling potential explanation of D'Asaro's (1995) observation that the NIW frequency shift was at least five times smaller than $\frac{1}{2}\zeta$ during Ocean Storms.

This interpretation of the Ocean Storms Experiment was pursued further by Balmforth et al. (1998), who ran spin-down simulations of the YBJ equation. NIWs were initialized in the mixed layer and evolved in the presence of an idealized, barotropic mesoscale eddy field. Qualitative comparisons between the simulations and observations showed that YBJ dynamics were not inconsistent with the observed time for NIWs to escape the mixed layer. Balmforth and Young (1999) showed that including the β-effect improved the agreement with observations. Because an idealized eddy field was used, however, no quantitative conclusions could be drawn about the ability of YBJ to capture the observed evolution.

More recently, Asselin and Young (2020) investigated the fate of NIWs as they propagate into a baroclinic mesoscale eddy field using numerical simulations of the extended YBJ system that also accounts for the effect of the waves on the mesoscale. They observed the initially horizontally uniform NIWs undergoing scale reduction by ζ-refraction and then propagate downwards in anticyclones. The vertical wave propagation terminated at depth due to the decay of the baroclinic vorticity away from the surface. For strong NIWs, they also found that the mesoscale eddy field was weakened due to stimulated generation. While this work was motivated by observations, it again employed an idealized setup that made direct comparisons to observations difficult.

In this study, we aim to bridge the gap between theory and observations by using the YBJ framework to interpret the observed evolution of NIWs in the Northeast Atlantic Ocean. We use an array of nine moorings to capture some of the mesoscale variations in the NIW field. The YBJ framework allows us to attribute the observed NIW evolution to a set of well-defined physical processes. We integrate the three-dimensional YBJ equation using observational inputs for the wind forcing, mesoscale streamfunction, and stratification, and we compare these simulations to simpler slab models. We show that the YBJ model offers significant improvements in modeling NIW evolution, without the need for any tuning. We use the YBJ energy budgets to provide a dynamical interpretation of spatial and temporal variations in the NIW field and quantify the relative importance of the various physical processes involved. We also provide an estimate for the importance of stimulated generation in this region.

2.3 Observations

2.3.1 NIW Data

We extract observations of NIWs from data collected for the Ocean Surface Mixing, Ocean Submesoscale Interaction Study (OSMOSIS; Buckingham et al., 2016). As part of the study, measurements were taken from nine moorings anchored over the Porcupine Abyssal Plain in the northeastern part of the Atlantic Ocean from September 2012 to September 2013. The moorings were all distributed less than 10 km from a central mooring located at (48.6875°N, −16.1875°E) (Fig. 2.1a,b). In this study, we focus on observations from the central mooring and the four outermost moorings. Each mooring took measurements of the horizontal velocity once every ten minutes using an acoustic current meter (ACM). On the central mooring, there were 13 ACMs spaced nominally between 50 m and 500 m depth. On the outer moorings there were five ACMs nominally spaced over the same depth range. The deepest ACM failed on the outer northeast (ONE) mooring, and so measurements are only available to a depth of 361 m.

The NIW signal is extracted using a Gaussian filter in frequency space. At this latitude, the near-inertial frequency (0.40 h^{-1}) is close to the frequency of the M_2 tidal constituent (0.50 h^{-1}), and so we choose a filter width which corresponds to 11% of f to exclude the

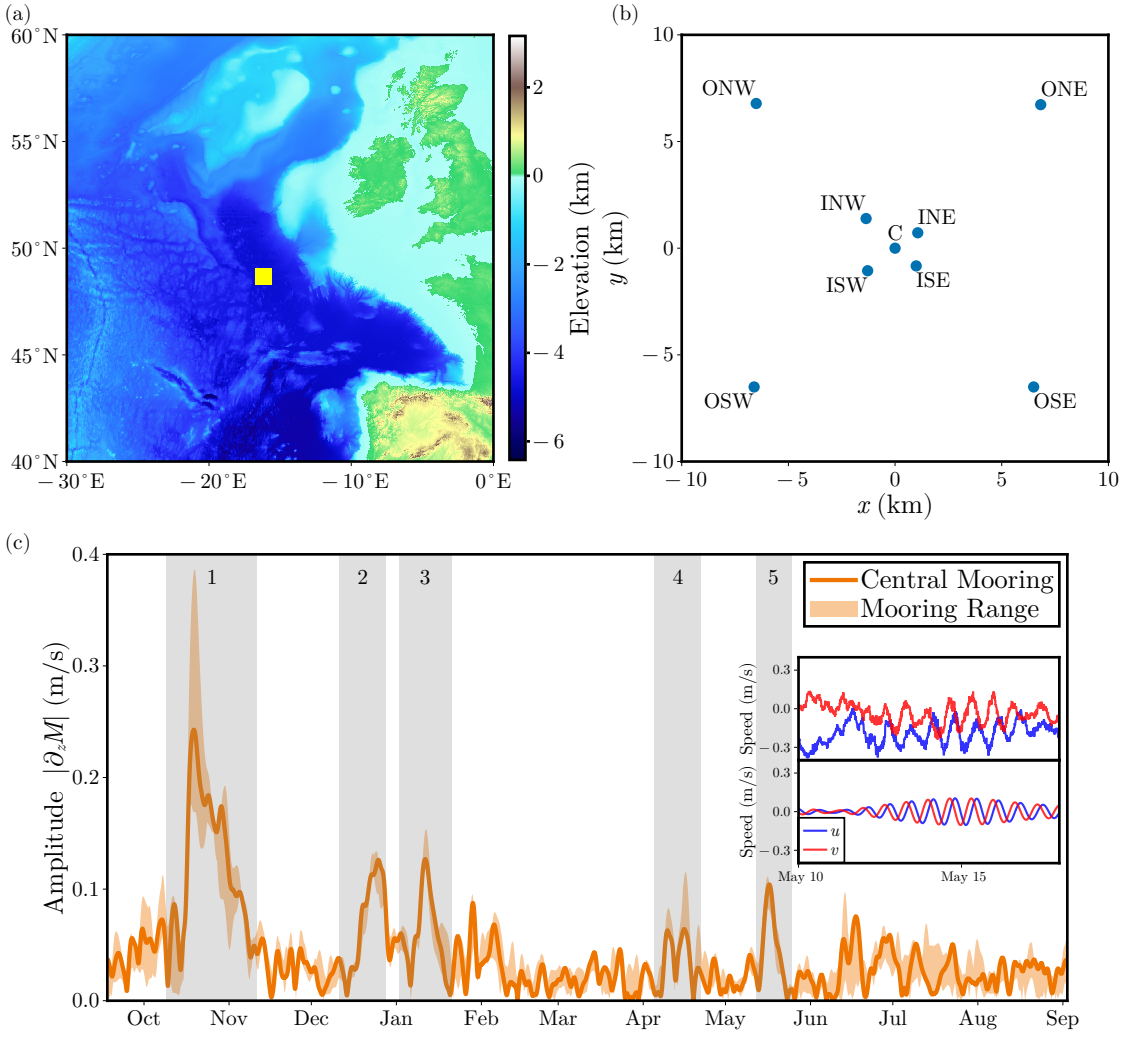

Figure 2.1: (a) Location of the OSMOSIS mooring array (yellow square) overlaid on a map of the bathymetry. (b) Geometry of the mooring array that consists of a central mooring (C) surrounded by an inner and an outer group of moorings. The inner moorings are labeled with an "I" followed by the cardinal direction of the mooring in relation to the central mooring. The outer moorings are labeled in a similar way but with an "O" identifier. (c) Time-series of NIW amplitude extracted from the OSMOSIS mooring observations from September 2012 to September 2013. The solid line indicates the observations at the central mooring while the orange ribbon represents the range across the mooring array. The gray shading indicates the five NIW events discussed in the text. To illustrate the process of extracting the NIW signal, the inset shows the raw velocity (top) and NIW velocity (bottom) during event 5. The amplitude of the envelope modulating the NIW velocity is $|\partial_z M|$.

sizable tidal motion from the filtered time series. The conclusions we draw below are not sensitive to the filter width so long as the filter does not include the M_2 peak. We identify five events in the year-long time series where strong NIWs were forced relatively coherently across the mooring array (Fig. 2.1c).

The actual depth of the instruments varied in time due to drag exerted by the flow on the moorings. Over the course of the year, there were several times where the moorings ex-

perienced knockdown by up to 200 m (Callies et al., 2020). We flag these events if the knockdown on the central mooring is more than 10 m when averaged with a 1-day running mean. Outside of these knockdown events the variations in mooring depth about the nominal depth is ∼2 m. The interpretation of the observed NIW signal during these knockdown events is complicated because the filter we use to extract the NIW signal does not commute with evaluating the signal along the trajectory of the ACMs. We do not attempt to explain the observations during the knockdown periods.

2.3.2 Stratification and mixed layer Depth

We need the stratification and mixed layer depth as observational inputs to the numerical YBJ simulations. The stratification influences the dispersion term in the YBJ equation and the mixed layer depth informs us about what depth range to force the waves over. Two ocean gliders sampled across the mooring array during the OSMOSIS study period (Damerell et al., 2016; Thompson et al., 2016). The gliders measured temperature and salinity in the upper 200 m of the water column approximately once every two hours. Following Damerell et al. (2020), we calculate the mixed layer depth H_m from the glider observations as the depth at which either the potential density ρ or the temperature T has changed more than a given threshold from the value at 5 m depth. These thresholds are $\Delta\rho = 0.03$ kg m^{-3} for density and $\Delta T = 0.2$ K for temperature. The final mixed layer depth is taken as the shallowest of the two estimates.

The stratification N^2 is calculated using:

$$N^2 = -\frac{g}{\rho}\frac{\partial \rho}{\partial z}, \tag{2.7}$$

where g is the acceleration due to gravity. The stratification is then averaged over a given event and the glider trajectories to result in a single spatially and temporally averaged profile used for simulations.

2.3.3 Wind Data

For the wind forcing, we use the European Centre for Medium-Range Weather Forecasting ERA-5 reanalysis (Hersbach et al., 2018). We extract time series of the 10 m zonal (u_w) and meridional (v_w) winds with hourly resolution at the grid point that contained the mooring array. Following Pollard and Millard Jr (1970) we convert this to a wind stress using a bulk aerodynamic drag formulation. In terms of the complexified wind velocity ($\mathcal{U}_w = u_w + iv_w$), the complexified wind stress τ_w is given by:

$$\tau_w = \rho_a C_D |\mathcal{U}_w| \mathcal{U}_w, \tag{2.8}$$

where ρ_a is the density of air and C_D is the drag coefficient, which we calculate using the speed-dependent formulation of Large and Pond (1981).

2.3.4 Altimetry

To characterize the mesoscale eddy field, we use observations of the sea surface height (SSH) from the Data Unification and Altimeter Combination System's (DUACS) delayed-time (DT) 2018 release (Taburet et al., 2019). The DUACS DT2018 SSH maps are provided at a nominal $(1/4)^\circ$ and daily resolution. We convert these measurements to a geostrophic streamfunction using $\psi = gh/f$, where h is the SSH and f is the latitude-dependent Coriolis parameter.

2.4 Models

2.4.1 The PM Model

We begin by writing the PM model in the language of YBJ. Adding (2.1a) to $i \times$ (2.1b) and multiplying by e^{-ift} yields

$$\partial_{tz} M = -r\partial_z M + \frac{\tilde{\tau}}{\rho_w H_m}, \tag{2.9}$$

where $\tilde{\tau} = e^{-ift}\tau_w$ is the back-rotated, complexified wind-stress. For each NIW event, we solve the PM equation using the reanalysis wind stress and the mixed layer depth from the gliders. In order to better understand the role of ζ-refraction in the life cycle of wind-forced NIWs, we perform a second run of the PM model (denoted by PM+ζ) augmented by the refraction term from the YBJ equation:

$$\partial_{tz} M = -r\partial_z M - \frac{i\zeta}{2}\partial_z M + \frac{\tilde{\tau}}{\rho_w H_m}. \tag{2.10}$$

This shifts the NIW phase and hence changes the phasing of the NIWs relative to the wind. It does not, however, capture the horizontal energy transfers induced by ζ-refraction because these transfers require horizontal structure in the wave field, which is absent from the horizontally uniform PM model.

For each event, we use a constant mixed layer depth that is an average of the time-varying mixed layer depth over duration of the event. Since the events are relatively short, the error in this approximation is minimal for most events. We make this choice of a constant mixed layer depth to avoid complications that arise otherwise, especially in the YBJ model discussed below.

The parameter r is intended to account for all of the processes that decrease the wave amplitude in the region of interest. This is primarily thought of as vertical propagation of NIWs out of the mixed layer, but other processes such as advection and dissipation may also cause the NIW amplitude to decrease. Treating all of these processes as a Rayleigh drag term with a single decay parameter represents a drastic simplification in the PM model. It also introduces a free parameter. Previous studies have chosen r such that the resulting solutions fit observations as closely as possible. The original PM paper used both $r^{-1} = 4$ days and $r^{-1} = 8$ days.

Alford (2001) used a damping that would correspond to about $r^{-1} = 4$ days at the latitude of the OSMOSIS mooring array. Yu et al. (2022a) used $r^{-1} = 16.7$ days to estimate the NIW wind work during the OSMOSIS experiment using the PM model. We speculate that they had to use very weak damping because their wind data was taken from ERA-interim reanalysis, which has 6-hourly analysis steps with forecasts used to increase the time resolution to 3 hours, which may suppress the wind power at frequencies important for NIW generation. A similarly weak damping is likely unsuitable here, given that ERA-5 reanalysis winds have more power in the near-inertial band (see discussion in Flexas et al., 2019). Nevertheless, we vary r^{-1} between 4 and 16 days.

For the PM+ζ model, we calculate the vorticity using the streamfunction from altimetry. The data processing is as described below for the YBJ model, and we select the vorticity value at the grid point nearest to the center of the mooring array.

We initialize the simulations with no waves and then allow the model to spin up before the main forcing period for each event. We choose the initial time to be when the observations show relatively little waves. This is done by eye. These initial times are followed by a strong forcing event and so the NIW signal is dominated by the newly generated waves, implying that the error from using an initial condition with no waves is relatively small. This discussion also applies to the YBJ simulations we run. We integrate both of the models (PM and PM+ζ) using a Crank–Nicolson scheme.

2.4.2 The YBJ Model

The YBJ equation is a three-dimensional partial differential equation, making it substantially more computationally expensive to integrate than the PM model. We solve the YBJ equation using the pseudospectral solver Dedalus (Burns et al., 2020) with a mixed explicit and implicit diagonal RK2 scheme. As discussed above, we start the simulations with no waves. We use a domain that is 400 km$\times$400 km in the horizontal (centered on the mooring array) and 4 km deep. Each dimension is discretized with 128 modes. The vertical dimension is finite and represented using Chebyshev polynomials. The horizontal dimensions are made periodic and represented using Fourier modes. The stratification and the wind forcing are taken to be horizontally uniform, capturing the forcing at a scale much larger than the mesoscale. Smaller-scale structure in the wind stress can generate smaller-scale NIWs, but the energy input tends to be strongly dominated by the large-scale winds (Rama et al., 2022b). To construct the mesoscale streamfunction used in the simulations, we take ψ from observations and calculate the vorticity by taking the finite-difference Laplacian of ψ on the sphere. We then interpolate this vorticity onto the Cartesian simulation grid centered on the location of the central mooring. We apply a taper to the vorticity, such that it goes to zero on the domain boundaries. We invert this tapered vorticity field for the streamfunction on the periodic domain, making the resulting streamfunction periodic as well. Our conclusions are neither sensitive to the form nor width of the taper. For the results shown below, we use

the following form of the taper:

$$\mathcal{T}(x,y) = \cos^2\frac{\pi x}{L_x}\cos^2\frac{\pi y}{L_y}, \tag{2.11}$$

where L_x and L_y are the x and y lengths of the domain, respectively.

The YBJ equation can be formulated on a β-plane by making the substitution $\zeta/2 \to \zeta/2 + \beta y$. We perform all of our analysis on an f-plane (i.e., $\beta = 0$) for two reasons: (i) We focus our analysis on a region in the Northeast Atlantic Ocean where $\beta L \ll \zeta/2$ with L a typical meridional scale of the waves. This relative scaling varies regionally in the ocean. For example, Thomas et al. (2020) studied a region that was similarly dominated by the vorticity and found that the ζ-refraction process was more important to the NIW evolution there than β-refraction, whereas β-refraction appeared to be important for Ocean Storms (D'Asaro et al., 1995). (ii) The β term adds difficulty in simulating the YBJ equation numerically (see Balmforth and Young, 1999).

For the forcing, we need to prescribe a vertical profile that determines at what depths the momentum is deposited. We specify the profile for the body force $\partial_z \mathcal{F}$ such that it is constant in the forcing layer and then decays rapidly to zero below:

$$\partial_z \mathcal{F} = \frac{\alpha\tilde{\tau}}{H_m\left[\alpha + \ln(2\cosh\alpha)\right]}\left[1 + \tanh\left(\alpha\frac{z+H_m}{H_m}\right)\right], \tag{2.12}$$

where α is a parameter determining the steepness with which the body force falls off below the forcing layer. We use $\alpha = 2$ in all the simulations presented below. The prefactor ensures that $\mathcal{F} = \tilde{\tau}$ at $z = 0$.

When the winds blow on the ocean surface, they generate turbulence that mixes momentum downwards. If a mixed layer already exists and is not too deep, the momentum input from the wind will be rapidly homogenized within the mixed layer (Pollard and Millard Jr, 1970; Kato and Phillips, 1969). For some events at the OSMOSIS site, the mixed layer was up to a few hundred meters deep. These deep mixed layers are likely the result of convection driven by buoyancy forcing rather than the mechanical wind forcing (Thompson et al., 2016). In these cases, it is unlikely that the momentum is uniform across the mixed layer, especially if the buoyancy forcing has ceased. The depth structure of NIWs obtained from the OSMOSIS mooring confirms this picture. The waves are initially forced over a layer that is thinner than the mixed layer before they propagate to depth. To avoid forcing over an unrealistically large depth, we cap the forcing layer at 80 m. This value is guided by the observations and represents an average depth over which waves are forced when the mixed layer is deep. We discuss possible ways to improve this representation below.

In order to solve the YBJ equation, we must further specify vertical boundary conditions. The requirement that the vertical velocity is zero on the top and bottom of the domain translates to the requirement that M be horizontally uniform at the top and bottom boundaries

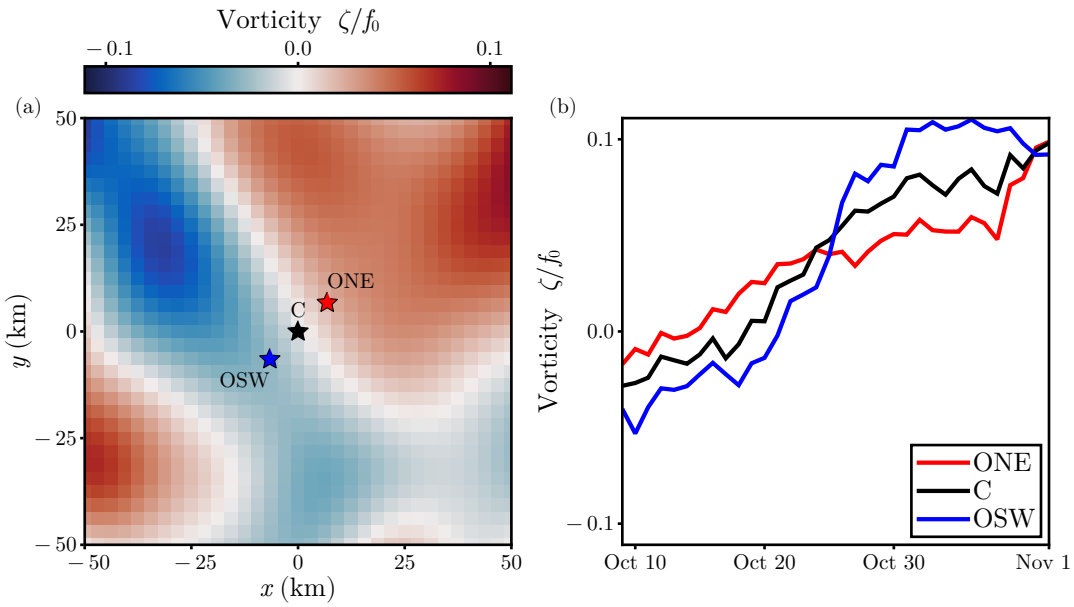

Figure 2.2: (a) Snapshot of mesoscale vorticity field in the inner quarter of the simulation domain on 2012-10-18 which was the peak of event 1. The stars show the locations of three specific moorings. The red star denotes the ONE mooring which is in a region of cyclonic vorticity, the black star denotes the central mooring and the blue star denotes the OSW mooring which is in an anti-cyclonic region. (b) Time series of the vorticity during event 1 at each of the three moorings denoted by the stars above. During the main forcing period the vorticity changes sign across the mooring array while at later times the entire mooring array transitions to being in a region of cyclonic vorticity.

(Young and Ben-Jelloul, 1997). Since M is determined by $\partial_z M$ up to an arbitrary horizontal function, without loss of generality we specify $M(x, y, -H, t) = 0$. The top boundary condition can be found by vertically integrating the YBJ equation twice and requiring that M be horizontally uniform at $z = 0$ (Appendix B). The result is

$$\partial_t M(x, y, 0, t) = \mathcal{F}, \tag{2.13}$$

which we integrate a priori and then use as a Dirichlet boundary condition on M.

2.5 Results

2.5.1 Case Study: Fall Event

We begin with a detailed analysis of the simulation results for event 1, which occurred in the fall (Fig. 2.1c). This event is by far the most energetic NIW event observed throughout the year. The main forcing for event 1 occurred when the mooring array straddled a dipole in the mesoscale vorticity (Fig. 2.2), making it a good candidate to see the effect of ζ-refraction.

NIW Amplitude and Phase

The observed NIW amplitude begins increasing around Oct. 14 and reaches a maximum value on Oct. 20 before returning to background levels by the end of our simulation period

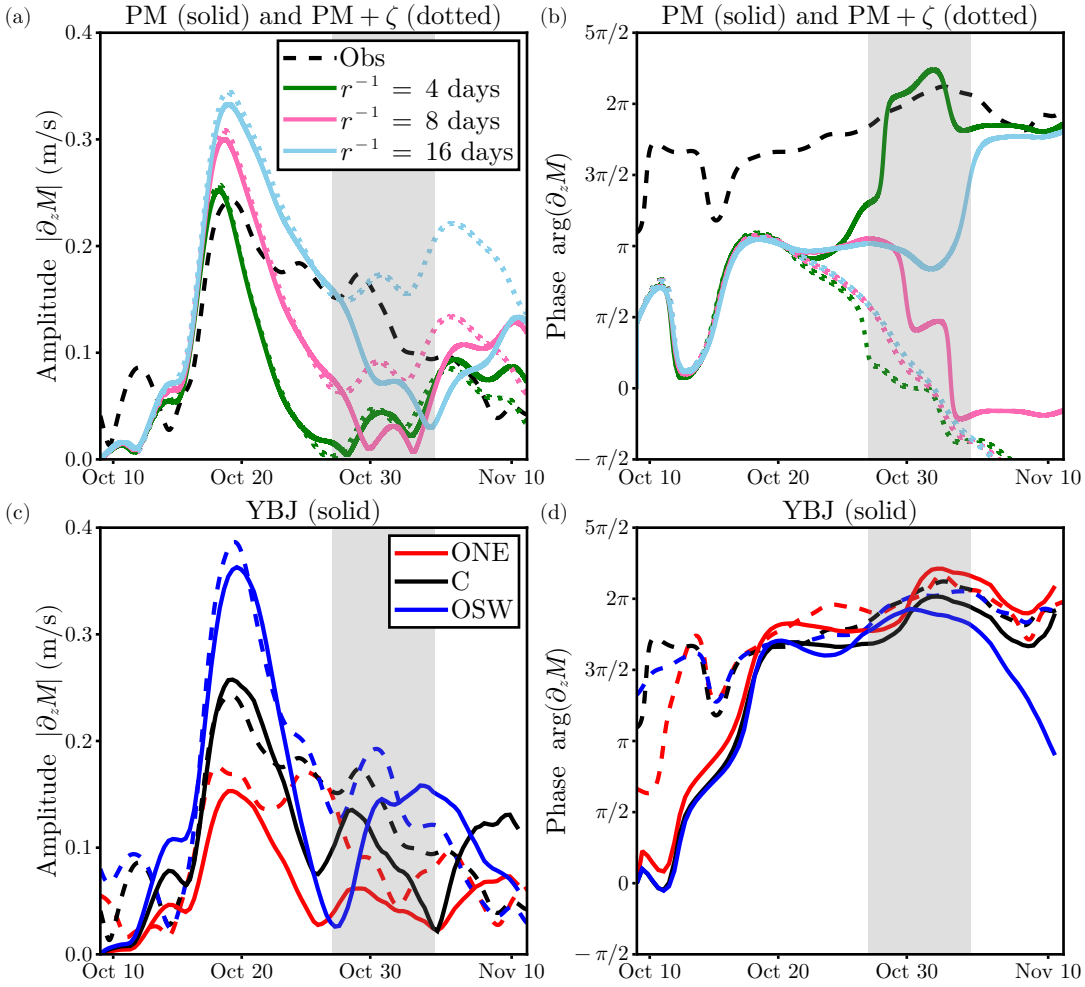

Figure 2.3: (a) NIW amplitude in the PM model compared to observations at the central mooring (dashed black line). The PM model was run with a range of difference values for the damping parameter r. Solid lines indicate the results of the PM model without refraction and dotted lines indicate the results with the refractive term added. The gray shaded region is a period of mooring knockdown. (b) As in (a) but for the NIW phase. (c) NIW amplitude in the YBJ simulation (solid lines) compared to observations (dashed lines) at the central mooring (black) as well as the ONE (red) and OSW (blue) moorings. (d) As in (c) but for the NIW phase.

on Nov. 10 (Fig. 2.3a). There are large variations in the peak amplitude over the mooring array, although these differences disappear by the end of the event (Fig. 2.3c). The phase increases sharply near the beginning of the forcing period as it aligns with the wind (Fig. 2.3b). Following this, it slowly increases for most of the event and is relatively homogeneous across the mooring array except for a period following the initial forcing (Fig. 2.3d).

The PM model has trouble capturing the NIW evolution during this event (Fig. 2.3a,b). Using a value of the damping parameter $r^{-1} = 4$ days results in a peak amplitude close to that of the observed NIW peak amplitude in the central mooring, but such a relatively strong damping causes the NIW amplitude to drop off much too quickly compared to the observations.

A more realistic amplitude decay is achieved when weaker damping is applied (especially with $r^{-1} = 16$ days), but then the peak NIW amplitude is overestimated substantially. The simulated phase bears little resemblance to the observed phase, with the simulated phase being offset by up to $\pi/2$ during the initial forcing period, and the simulated phase remains close to constant around π as long as substantial amplitude remains, missing the gradual increase in the observed phase.

The addition of the refractive term does little to change the simulated NIW amplitude in the PM model during most of the event (Fig. 2.3a,b), indicating that the phase shifts introduced by refraction are unable to substantially change the relative alignment between the NIWs and the winds. At later times the effect of adding refraction is more pronounced in the runs with lower damping as there is more time for refraction to act on the waves before they are dissipated. Refraction seems to dominate the phase evolution at all three damping values, but the phase trends in the opposite direction from what is observed. The phase tendency is also larger in magnitude than what is observed.

While the PM model may be able to capture some qualitative features of the NIW observations, it seems clear that: (i) the physical processes that cause the NIWs amplitude to decay cannot be accurately captured using a simple linear drag formulation, and (ii) if refraction is important to NIW evolution, its effect is not simply to change the phase of the NIWs but must involve horizontal processes.

The YBJ model captures the observed amplitude and phase evolution much better than the PM model (Fig. 2.3c,d). At the central mooring, the YBJ simulation agrees with observations in terms of peak amplitude and decay timescale. Similarly, the phase evolution is much closer to observations than the PM results, in terms of both its value after the forcing and its trend afterward. Again, we emphasize that, unlike the PM models, there is no tunable parameter in the YBJ simulations.

The YBJ model also captures observed lateral variations of the NIW signal across the mooring array (Fig. 2.3c,d). At the OSW mooring, which at the start of the event is in a region of anti-cyclonic vorticity (Fig. 2.2), the YBJ simulation successfully predicts a substantial enhancement in the NIW amplitude compared to the central mooring. At the ONE mooring, which at the start of the event is in a cyclonic region (Fig. 2.2), the YBJ simulation successfully shows a reduction in the NIW amplitude compared to the central mooring. The YBJ model also captures that the NIW phase is much more uniform across the mooring array than the amplitude.

NIW kinetic energy Budget

The observations show a higher NIW amplitude at the mooring in the anticyclone than at the mooring in the cyclone (Fig. 2.3c,d). This is suggestive of ζ-refraction concentrating NIW kinetic energy in anticyclones. It is hard to draw this conclusion from the observations

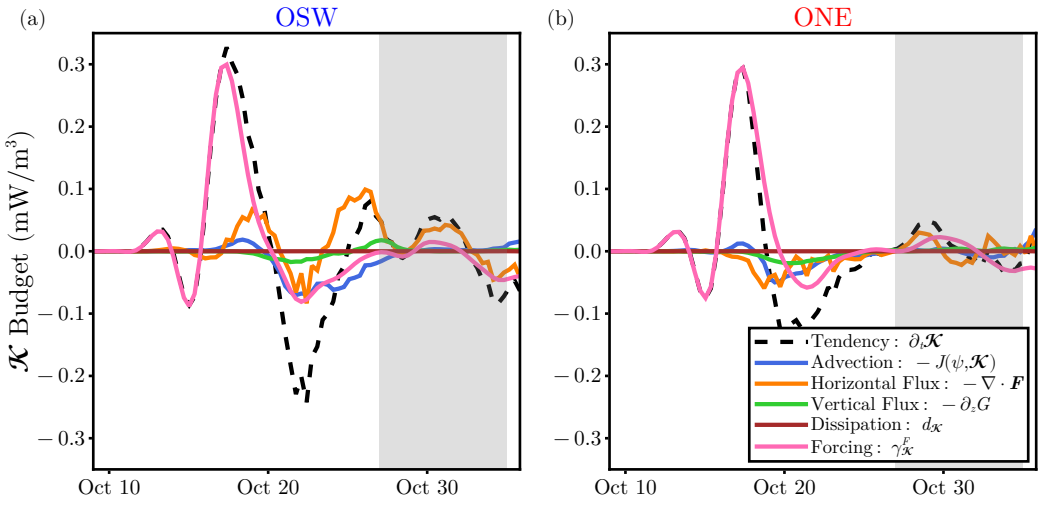

Figure 2.4: (a) NIW kinetic energy budget terms. The kinetic energy tendency (dashed line) is decomposed into the 5 processes in the model which can change the kinetic energy: advection (blue), horizontal flux divergence (orange), vertical flux divergence (green), hyperviscosity (brown) and wind forcing (pink). The budgets are evaluated at the horizontal position of the moorings and at fixed depth of 25 m. (b) As in (a) but for the ONE mooring. To better visualize the terms we only plot the budget for the first 3/4 of the event.

alone, however, because other factors could give rise to the observed amplitude differences. For example, the current meters were located at slightly different depths (between $z = -44$ m and $z = -62$ m).

With the YBJ simulations capturing the observed differences between the moorings, we can use the YBJ framework to identify the processes giving rise to these lateral variations. We construct a point-wise kinetic energy budget (2.4) at the OSW (anticyclonic region) and ONE (cyclonic region) moorings, which allows us to separate the kinetic energy tendency into advection, dispersive flux divergence, dissipation, and forcing (Fig. 2.4). During the initial forcing period (up to Oct. 20), the tendency due to wind forcing is similar for both moorings. At both locations, there is a small positive advective tendency that turns slightly negative toward the end of the forcing period. The vertical flux divergence is also similar at the two locations and smaller in magnitude than the advective tendency. Dissipation is negligible. The most notable difference between the two mooring locations is in the horizontal flux divergence term. There is horizontal flux convergence at the OSW mooring and flux divergence at the OSE mooring. This causes the total tendency to be larger than the wind forcing at the OSW mooring and smaller than the wind forcing at the ONE mooring.

At these early times, ζ-refraction is the primary driver of the horizontal energy flux and causes concentration of NIW kinetic energy into anticyclonic regions. During the peak of the event, the horizontal energy flux is directed from cyclonic to anticyclonic regions (Fig. 2.5a). This arises from an interplay of the refraction and dispersion terms in the YBJ dynamics: re-

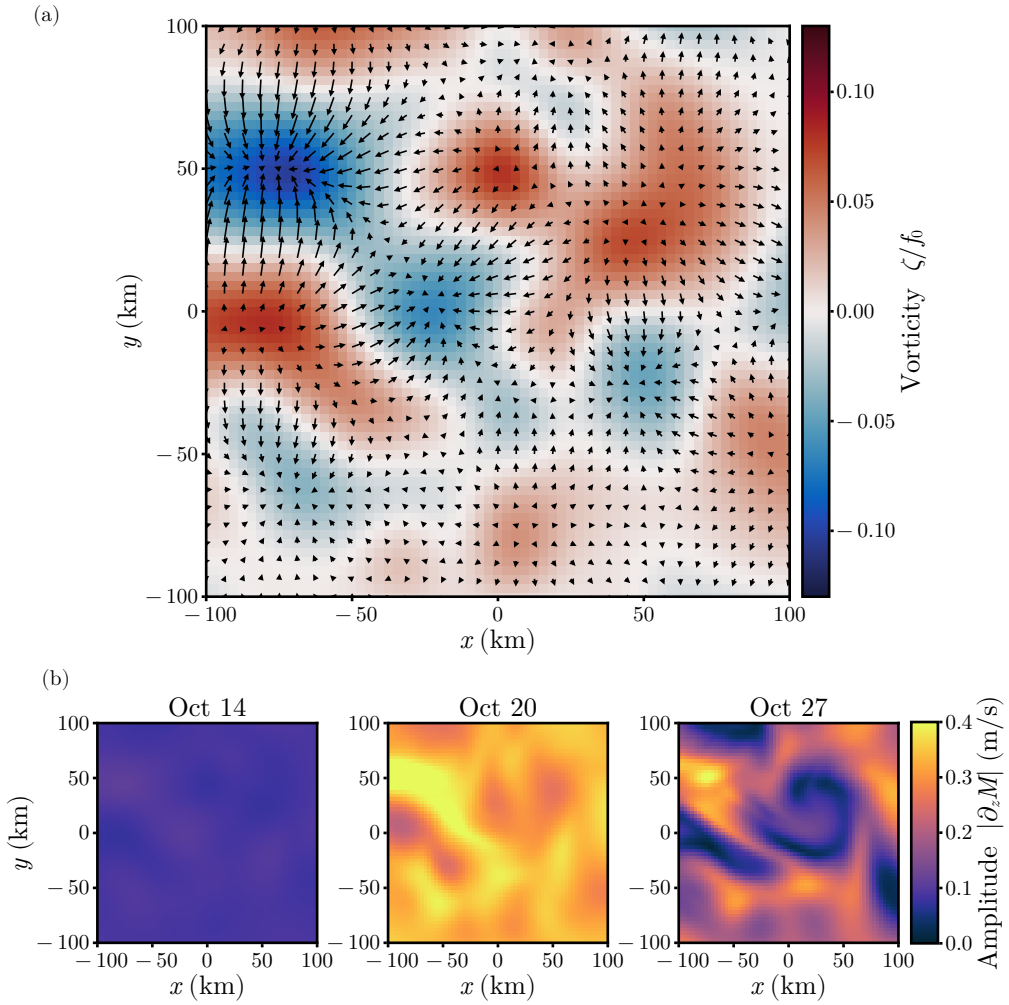

Figure 2.5: (a) Color map of vorticity (normalized by f) at the peak of event 1. Arrows indicate the horizontal energy flux field at $z = -25$ m, showing a flux of NIW kinetic energy out of cyclones and into anticyclones. (b) Horizontal sections of NIW amplitude at a depth of 25 m and three different times. Sections show the inner half of the simulation domain. There is no horizontal structure to the forcing and hence the NIWs are initially forced uniformly throughout the domain. Refraction is the only process which can impose structure on a uniform NIW field in the YBJ framework. NIWs begin to be concentrated into anticyclones. Brighter regions in the Oct. 20 plot correlate with anticyclonic regions in the altimetry. Once refraction creates horizontal structure other processes can act. Dispersion will eventually counteract concentration into anticyclones. Advection by the mesoscale eddies will also stir horizontal structure created by refraction. A signal of advective stirring is clearly visible in the upper right quadrant of the Oct. 27 plot.

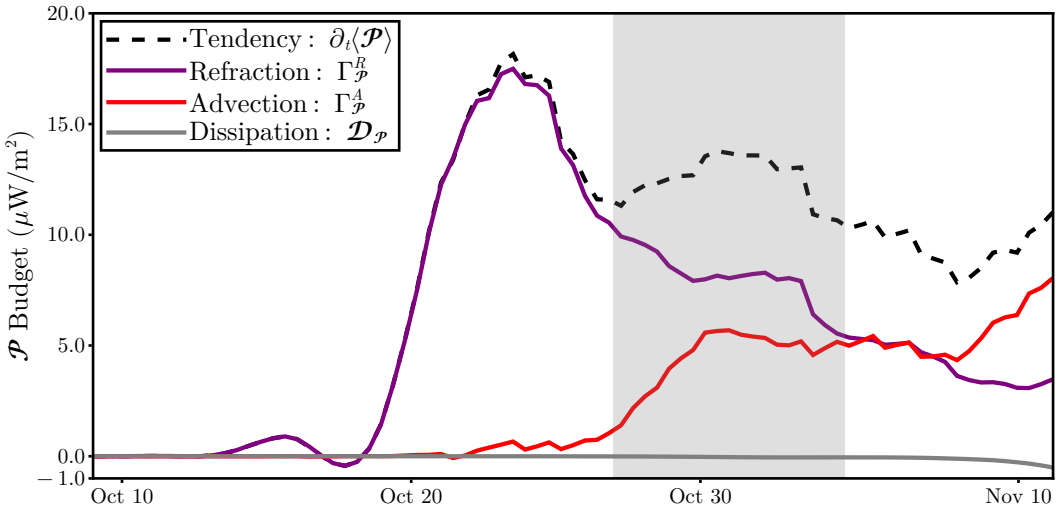

Figure 2.6: Vertically integrated and horizontally averaged potential energy budget showing the potential energy tendency (dashed line) split into contributions from refractive production (purple), advective production (red) and hyperviscosity (gray).

fraction sets up phase gradients, which cause a dispersive energy flux as described by (2.5). As horizontal structure develops, advection can also become important as it stirs the existing horizontal structure. This sequence of events was described by Rocha et al. (2018) and captures the early evolution during this event (Fig. 2.5b).

We again emphasize that the PM model, even if the refractive term is included, cannot capture these dynamics because lateral energy transport originating from the dispersion term in the YBJ equation is crucial. Once dispersion is included, the YBJ model captures observed lateral variations in the NIW amplitude between anticyclonic and cyclonic regions and offers a clear dynamical explanation.

NIW potential energy Budget

The NIW potential energy budget is of interest because it provides insight into the energy exchange with mesoscale eddies (Xie and Vanneste, 2015; Rocha et al., 2018). While we prescribe the mesoscale eddy field using altimetry rather than evolving a coupled system, we still interpret the sources of NIW potential energy as estimates of the energy transfer from mesoscale eddies facilitated by NIW refraction and advection.

The vertically integrated and domain-averaged budget for event 1 shows that NIW potential energy is generated by both refraction and advection (Fig. 2.6). The potential energy tendency rises sharply from zero to a peak value after the main forcing period of the event, lagging the peak in the kinetic energy tendency by a few days. At these early times, almost all of the NIW potential energy is created by refractive production. As horizontal structure is created, advective production ramps up. The refractive production term decreases through-

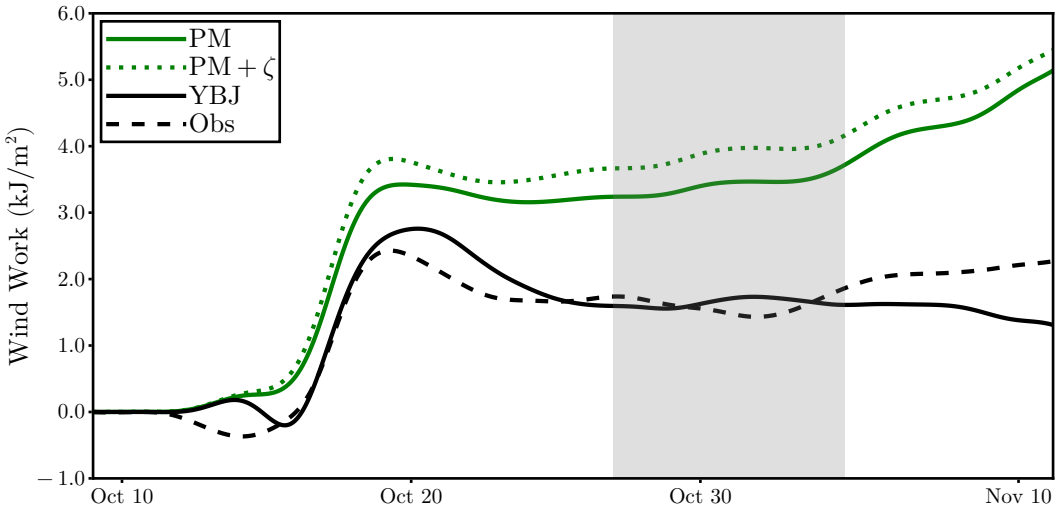

Figure 2.7: Wind-work as estimated from observations at the central mooring (dashed black), PM (solid green), PM+ζ(dotted green) and YBJ (solid black). For observations the wind-work is defined as $\tau_{NIW} \cdot \mathbf{u}_{NIW}$, while for the simulations it is defined as the kinetic energy production term from the respective kinetic energy budget. The PM models use $r^{-1} = 4$ days which best matches the peak amplitude.

out the rest of the event and by the end is overwhelmed by advective production. The initial dominance of refractive production followed by an increase in advective production is very similar to the succession of events Rocha et al. (2018) described for stimulated generation in idealized simulations that included the full coupling with the mesoscale dynamics.

The potential energy production—and presumed sink of mesoscale energy—peaks at about 20 μW m^{-2}. To put this number in context, a global input of 1 TW into the mesoscale eddy field (Wunsch and Ferrari, 2004) corresponds to about 3 mW m^{-2} on average. For stimulated generation to be important in the global energy budget of mesoscale eddies, it must be much stronger elsewhere.

NIW Wind Work

The interaction between NIWs and mesoscale eddies also affects the wind work in the near-inertial band (Fig. 2.7). The PM and PM+ζ models with $r^{-1} = 4$ days best match the peak amplitude in the NIW evolution (Fig. 2.3a) but overestimate the wind work at the location of the central mooring by a factor of more than two. We could also tune r to match the wind work estimated directly from observations integrated over the event. This is achieved with $r^{-1} = 0.576$ days, but this means that the peak NIW amplitude is underestimated by a factor of three and violates $r \ll f$. The effect of refraction in the PM model is to slightly increase the wind work. In the YBJ model, in contrast, the wind work matches the observations well. This is a consequence of the YBJ model's ability to closely reproduce the observed NIW evolution. While differences in the wind work between the YBJ and PM models appear

substantial, we stress that this event is unlikely to be representative of a time and space average. We intend to discuss this difference further in a subsequent publication.

2.5.2 Further Events in the Time Series

Simulations of the remainder of the NIW events (Fig. 2.8) confirm that the YBJ model better captures the observed NIW amplitude and phase evolution than the PM model does (Fig. 2.8). Event 4 is different from the others in that it consists of a double peak in NIW amplitude without a decay to background levels in between. The mixed layer depth was much more variable than during other events, especially towards the end of the event. For that reason, we simulate event 4 in two parts (events 4a and 4b) in order to minimize the variations in the mixed layer depth over a simulation period.

In general, we see that the YBJ model performs better than the slab models in reproducing the observed evolution of the NIW amplitude and phase. In event 2, the YBJ model captures the slow rise of the NIW amplitude as well as the decrease at late times. The PM models are not able to capture this behavior as well (Fig. 2.8a). All of the models, however, have substantial errors in the phase at early times. This could be due to pre-existing NIWs that we do not capture because we initialize the models at rest. The notable exception to the trend that YBJ performs better than the PM models is event 3 (Fig. 2.8b), where the YBJ model not only overestimates the peak but also the decay time of the waves. YBJ does, however, capture the slow evolution of the phase better than the PM models. We discuss some of the potential reasons for these disagreements below. For event 4a, all models predict a double peak in the amplitude that is not seen in observations (Fig. 2.8c). The YBJ model does better at later times in both amplitude and phase. In event 4b, the YBJ model does rather well in predicting the complete evolution of both the NIW amplitude and phase (Fig. 2.8d). The YBJ model predicts the timing of the peak better than the PM models in event 5, but the waves persist for longer than observed (Fig. 2.8e).

2.6 Discussion

The YBJ model does well in reproducing the mooring observations of NIWs. There are still differences between the model and the observations, however, as well as a couple of events where the YBJ model does less well. Given the observational inputs to the model, it would be surprising if this were not the case. One major limitation is the use of altimetry for the mesoscale streamfunction. As mentioned previously, the altimetry data are a heavily smoothed version of the real field. We suggested above that some of the discrepancies between the YBJ model and the observations were due to this smoothing. We also interpolated the mesoscale vorticity onto a smaller simulation grid. The resolution of the altimetry product is $0.25° \times 0.25°$, which is larger than the size of the mooring array. Differences across the mooring array come from the interpolation between neighboring altimetry grid cells, which will have introduced interpolation errors.

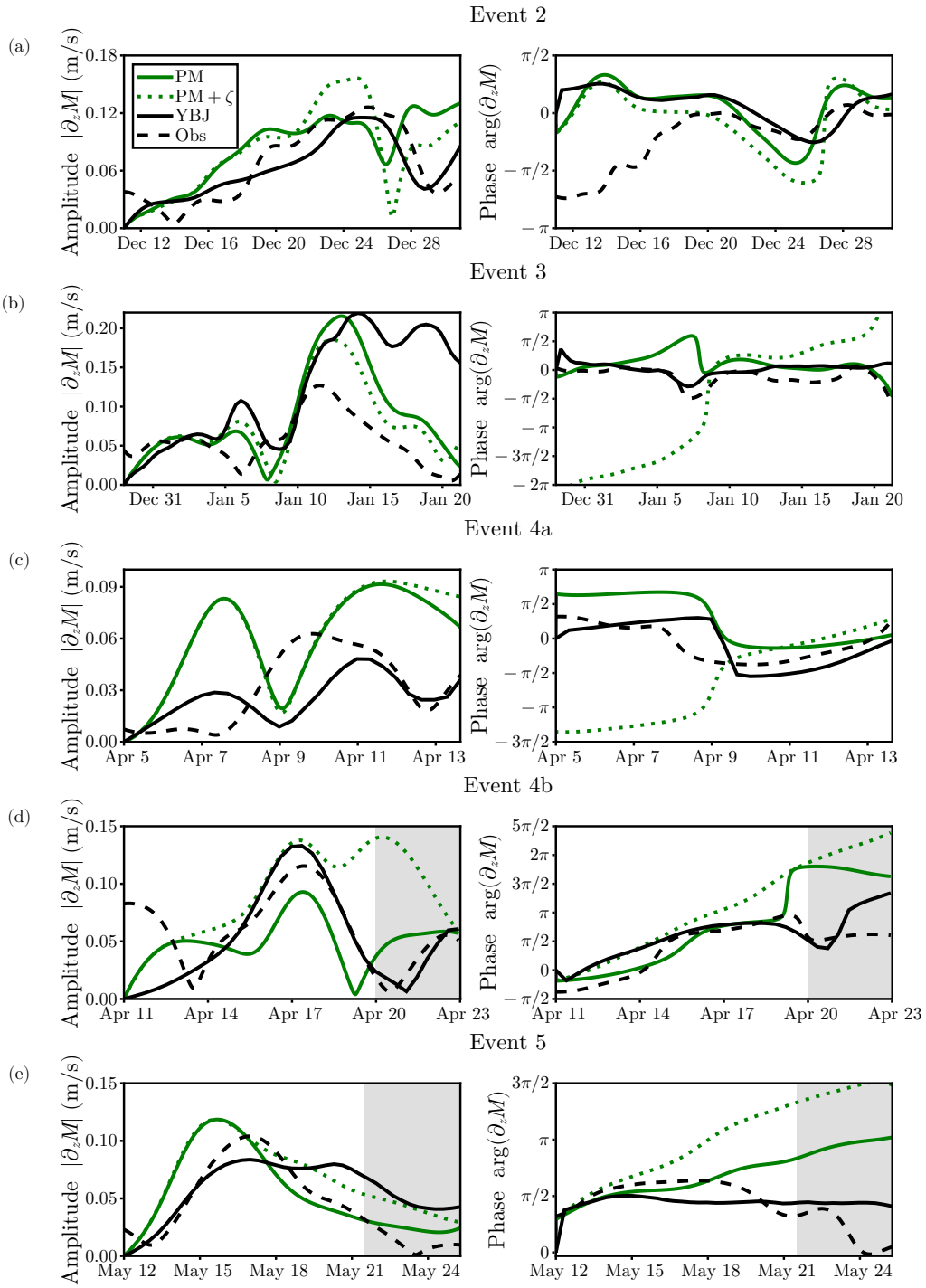

Figure 2.8: NIW amplitude (left column) and NIW phase (right column) in observations at the central mooring (dashed black line) as compared to PM (solid green) PM+ζ (dotted green) and YBJ (solid black) for (a) event 2, (b) event 3, (c) event 4a, (d) event 4b and (e) event 5.

Furthermore, we assumed that the mesoscale eddy field was barotropic. This is a reasonable assumption if the vertical scale of the waves is much smaller than the vertical scale of the eddies. There are certainly errors in the YBJ evolution, however, that arise from neglecting the baroclinicity of mesoscale eddies. These effects could be investigated in the future by running similar simulations using an *in-situ* data set that resolves the vertical structure of one or more eddies.

In event 1, the vorticity has large variations across the mooring array (Fig. 2.2a). This is reflected in the NIW observations as large differences in the amplitude across the mooring array. However, this need not be the case for all events. Event 3, by contrast, shows weak variations in the vorticity (as diagnosed from altimetry) across the mooring array. The result is that the YBJ simulations also show weak variations in the NIW field across the mooring array region. The altimetry vorticity field is a heavily smoothed representation of the real vorticity field of the ocean; smaller scale vorticity features are invisible to measurements from altimetry. If there is little variation in the larger-scale vorticity across the mooring array, then these smaller-scale vorticity features may play a more important role in governing the variations in $\partial_z M$. While we do not have a spatial map of smaller-scale vorticity features, we can estimate their magnitude at the mooring array by calculating the vorticity by applying Stokes' theorem to the area bounded by the outer moorings (Buckingham et al., 2016). The velocity that we use in Stokes' theorem is low-pass filtered to estimate the balanced signal. This vorticity agrees in general with the vorticity calculated from SSH but shows more high-frequency variability (Fig. 2.9). Event 3 shows the biggest disagreement between the two estimates of all the events. The mooring estimate shows cyclonic vorticity, whereas the altimetry estimate shows anticyclonic vorticity. This likely explains why we see the amplitude decay much quicker in observations compared to simulations because there would have been a horizontal energy flux out of the region while the simulations have a flux into the region, which acts to maintain the amplitude. The other events show better agreement between the two vorticity estimates, although there are times where the deviation is larger. Events 1 and 2 specifically show two anticyclonic periods that are not captured by altimetry. These may explain some of the mismatch between simulations and observations. For example, the anticyclonic excursion in event 1 probably explains why the YBJ simulation results decay more quickly than observations. For event 2, the excursion occurs primarily near the start of the event, where the NIW amplitude is weak, so its effect is tempered. Many of these anomalies are short in duration, which limits the error in using the altimetric vorticity.

The vorticity error for event 4a is minimal (Figure 2.9). We suspect the mismatch between the YBJ model and the observations in this event is due to the forcing. The reanalysis product used is not the exact wind-forcing felt at the OSMOSIS site. Both YBJ and PM show an initial peak in the NIW amplitude that is not seen in the observations, indicating that the wind-stress may be wrong at this point. The forcing event seen in observations then has to destroy these waves before forcing new ones, which causes the lower amplitude of the NIW

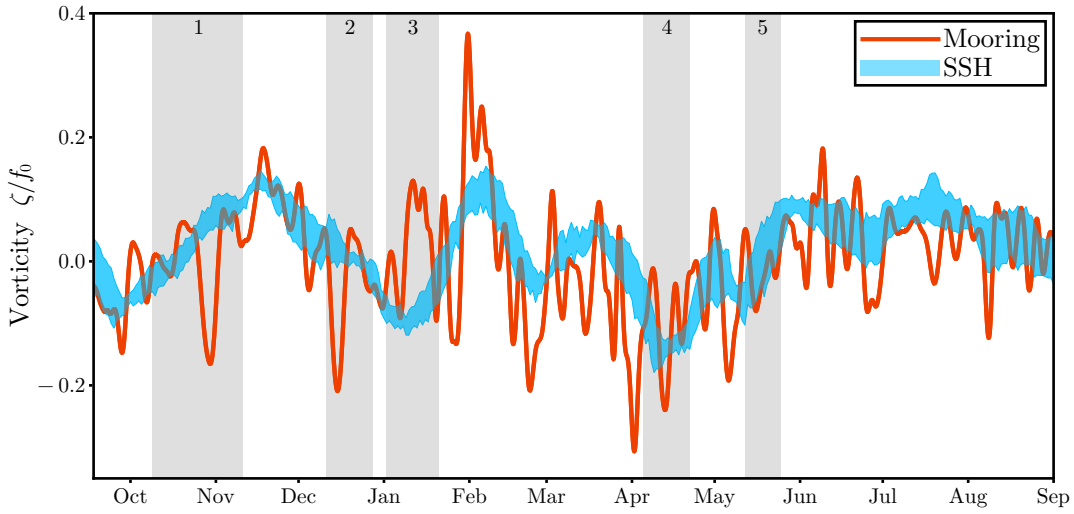

Figure 2.9: Comparison of the vorticity as calculated from Stokes' theorem applied to the outer moorings (orange) and from altimetry (blue). The velocities used in the Stokes' theorem calculation were obtained by low-pass filtering with a second order Butterworth filter and a cutoff frequency corresponding to a period of approximately 5 days. The blue ribbon shows the spread in the vorticity when interpolated onto the outer mooring positions.

peak in the YBJ simulation compared to observations.

Notwithstanding the caveats above, it is significant that the YBJ model can reproduce much of the NIW evolution with only the mesoscale vorticity as derived from altimetry. One may expect the larger-magnitude submesoscale vorticity to be at least as important for the NIW evolution. But the dispersion term in the YBJ equation depends on the Laplacian of M and hence in spectral space scales as κ^2. This means that the refractive generation of small-scale structure in the wave field will be opposed by increasingly strong dispersion. Our results therefore suggest that, at the location of the mooring array, dispersion indeed outpaces refraction at submesoscales and mesoscale refraction is more important for the NIW evolution (cf., Yu et al., 2022a).

Using a time-averaged stratification profile likely also contributes to differences between the observations and simulations, despite our attempts to keep simulation times as short as possible to avoid this effect. The assumption that the stratification does not vary in time is inherent in the YBJ scaling assumptions. Related to this is the question of what depth to force the model over. We used a fixed forcing depth, although this certainly varied over the simulation periods. This problem could be side-stepped by representing the wind forcing as a surface stress and adding a parameterized turbulent vertical momentum flux to the YBJ equation to transfer momentum from the wind downwards.

The YBJ equation contains no term to represent the breaking of NIWs. The scaling $H^* = u^*/\sqrt{N^* f}$, where $u^* = \sqrt{\tau/\rho_0}$ is the friction velocity and N^* is the stratification at the base of the mixed layer, indicates the depth to which wind-driven turbulence could deepen the

mixed layer (Pollard et al., 1973). If the mixed layer depth is shallower than H^*, the effect of NIW dissipation would be important as there is potential for mixed layer deepening by breaking of NIWs (Plueddemann and Farrar, 2006). During the OSMOSIS study, Yu et al. (2022a) showed that the mixed layer was generally deeper than H^*. This is especially true outside of summer and during the events we consider here. This is due to strong surface buoyancy forcing deepening the mixed layer by convection to a depth beyond what the winds alone could do. While one could include a critical Richardson number criterion to add NIW breaking to the YBJ equation, we believe that it is justifiable to ignore dissipation of NIWs in the surface layer in this study.

In event 1, the wind work predicted by YBJ agreed well with that calculated by observations. The PM model greatly overestimated the wind-work. The YBJ and PM models also disagree on the wind work for all the other events (not shown). If mesoscale interactions change the alignment of the NIWs with the winds, then this results in a very different wind work. The difference in wind work between the YBJ and PM models depends on the degree to which the mesoscale changes the alignment of the waves with the wind and the power of the winds at different frequencies. The differences in wind work can be large for individual events but the five events we simulated are not enough to estimate what the average difference might be when averaged over many events. Nonetheless, this may provide motivation to revisit previous estimates of the NIW wind work in order to determine whether the mesoscale may modulate wind work into the NIW band globally.

2.7 Conclusion

The evolution of the NIW field at the OSMOSIS site in the Northeast Atlantic Ocean is strongly modulated by mesoscale eddies. The observed evolution can be captured by the YBJ model, which includes NIW refraction, advection, and dispersion. If these processes are omitted as in the PM model, the observations cannot be reproduced as well, even if the parameterized damping rate is tuned.

The YBJ model provides a powerful interpretive framework for understanding observations of NIWs in the upper ocean. It allows us to attribute the observed evolution to specific physical processes. Lateral differences in the NIW amplitude across the mooring array are caused by ζ-refraction, which causes NIW kinetic energy to be fluxed into anticyclonic regions. While correlations between NIW amplitudes and mesoscale vorticity can be diagnosed from the observations alone (e.g., Yu et al., 2022a), the YBJ framework produces a quantitative prediction for the NIW amplitudes that matches observations and provides a physical interpretation.

The YBJ model also allows us to calculate changes in the NIW potential energy, which are expected to arise from transfers of balanced mesoscale energy (Xie and Vanneste, 2015; Rocha et al., 2018). For the strongest NIW event observed during the OSMOSIS campaign,

the NIW potential energy gain is at least two orders of magnitude smaller than the global average energy input into mesoscale kinetic energy. Stimulated generation therefore does not appear to have a major impact on the mesoscale eddy field in this part of the ocean.

Acknowledgments The authors thank two anonymous reviewers whose comments helped improve the presentation of this work. This material is based upon work supported by the National Science Foundation under Grant No. OCE-1924354 and by the National Aeronautics and Space Administration under Grant No. 80NSSC22K1445 issued through the Science Mission Directorate (Future Investigators in NASA Earth and Space Science and Technology).

Data Availability Statement The code to run the 3D YBJ model is available at `https://github.com/scott-conn/3DYBJ`. All of the OSMOSIS data used as part of this study is available from the British Oceanographic Data Centre. The OSMOSIS mooring data is available at `https://www.bodc.ac.uk/data/bodc_database/nodb/data_collection/6093/`. The OSMOSIS glider data is available at `https://doi.org/10.5285/6cf0b33e-a192-549f-e053-6c86abc01204`. The SSH data is available from the E.U.'s Copernicus Marine Service at `https://doi.org/10.48670/moi-00148`. The ERA5 reanalysis data is available from the Copernicus Climate Change Service (C3S) Climate Data Store at `https://doi.org/0.24381/cds.adbb2d47`.

2.8 Appendix A: YBJ kinetic energy Budget

The YBJ kinetic energy budget may be formed by multiplying (2.2) by $-M^*/2$ and adding the complex conjugate. Using integration by parts, the resulting kinetic energy budget can be written as

$$\partial_t \mathcal{K} + J(\psi, \mathcal{K}) + \nabla \cdot \boldsymbol{F}'_{\mathcal{K}} + \partial_z G'_{\mathcal{K}} = d_{\mathcal{K}} + \gamma^F_{\mathcal{K}}, \tag{2.14}$$

where

$$\mathcal{K} = \frac{1}{2}|\partial_z M|^2, \tag{2.15a}$$

$$\boldsymbol{F}'_{\mathcal{K}} = \frac{iN^2}{4f}\left(M\nabla M^* - M^*\nabla M\right), \tag{2.15b}$$

$$G'_{\mathcal{K}} = -\frac{M^*}{2f}\left(\partial_{zt} M + J(\psi, \partial_z M) + \frac{i\zeta}{2}\partial_z M + \nu\nabla^4\partial_z M - \partial_z \mathcal{F}\right) + \text{ c.c.}, \tag{2.15c}$$

$$d_{\mathcal{K}} = \frac{\nu}{2}\left(\partial_z M^*\nabla^4\partial_z M + \partial_z M\nabla^4\partial_z M^*\right), \tag{2.15d}$$

$$\gamma^F_{\mathcal{K}} = \frac{1}{2}\left(\partial_z M^*\partial_z \mathcal{F} + \partial_z M\partial_z \mathcal{F}^*\right). \tag{2.15e}$$

In this form, however, the vertical flux does not vanish on the boundaries at $z = 0$ and $z = -H$. In order to meaningfully separate horizontal and vertical fluxes, we redefine the 3D kinetic energy flux vector $\mathbf{H}' = [\mathbf{F}', G']^{\mathrm{T}}$, taking advantage of the fact that only its divergence appears in (2.14). The kinetic energy budget is thus invariant under the transformation $\mathbf{H}' \to \mathbf{H} = \mathbf{H}' + \nabla_3 \times \boldsymbol{\chi}$ with some vector field $\boldsymbol{\chi}$. We propose to pick a $\boldsymbol{\chi}$ such that the

transformed G is zero on both boundaries. To determine χ, we first integrate (2.2) from $-H$ to z and then introduce a new field A such that $\partial_z A = N^2 M / f^2$, which allows us to evaluate the integral of the dispersion term. The resulting equation is

$$\partial_{zt} M + J(\psi, \partial_z M) + \frac{if}{2}\nabla^2 A + \frac{i\zeta}{2}\partial_z M - \partial_z \mathcal{F} + \nu\nabla^4 \partial_z M = \mathcal{C} \tag{2.16}$$

with

$$\mathcal{C}(x, y, t) = \partial_{zt} M + J(\psi, \partial_z M) + \frac{if}{2}\nabla^2 A + \frac{i\zeta}{2}\partial_z M - \partial_z \mathcal{F} + \nu\nabla^4 \partial_z M \quad \text{at} \quad z = -H. \tag{2.17}$$

We are free to choose $\mathcal{C}$ as a boundary condition on the new field A, i.e., we can choose A at $z = -H$ as the solution of (2.17) for some specified $\mathcal{C}$. A sensible choice for the transformation is thus setting $\mathcal{C} = 0$ and

$$\chi = \frac{if}{4}\begin{pmatrix} M^*\partial_y A - M\partial_y A^* \\ -M^*\partial_x A + M\partial_x A^* \\ 0 \end{pmatrix}. \tag{2.18}$$

Under this transformation, the kinetic energy budget becomes (2.4), with the horizontal and vertical fluxes given by

$$\mathbf{F}_{\mathcal{K}} = \frac{if}{4}\left(\partial_z M^* \nabla A - \partial_z M \nabla A^*\right), \tag{2.19a}$$

$$G_{\mathcal{K}} = \frac{if}{4}\left(\nabla M \cdot \nabla A^* - \nabla M^* \cdot \nabla A\right). \tag{2.19b}$$

It is clear that the vertical flux $G_{\mathcal{K}}$ vanishes at the boundaries, as desired, because $\nabla M = 0$ there.

The divergence of the flux terms vanishes when the budget is integrated over the entire domain, and the only terms that can alter the domain-integrated kinetic energy are the wind work and dissipation:

$$\partial_t \langle \mathcal{K} \rangle = \mathcal{D}_{\mathcal{K}} + \Gamma^F_{\mathcal{K}}, \tag{2.20}$$

where $\mathcal{D}_{\mathcal{K}} = \langle d_{\mathcal{K}} \rangle$ and $\Gamma^F_{\mathcal{K}} = \langle \gamma^F_{\mathcal{K}} \rangle$.

2.9 Appendix B: YBJ Upper Boundary Condition

Beginning from (2.16) with the choice $\mathcal{C} = 0$, we can vertically integrate from $z = -H$ to $z = 0$ and use $M = 0$ at $z = -H$ to arrive at

$$\left[\partial_t M + J(\psi, M) + \frac{i\zeta}{2} M - \mathcal{F} + \nu\nabla^4 M\right]_{z=0} = -\frac{if}{2}\nabla^2 \int_{-H}^{0} A \, dz. \tag{2.21}$$

The no-normal flow condition is imposed by requiring $\nabla M = 0$ at $z = 0$ (Young and Ben-Jelloul, 1997), which eliminates the advection and dissipation terms. We then horizontally average (denoted by $\overline{\cdot}$) equation (2.21). Because $M(x, y, 0, t)$ has no horizontal structure,

it is equal to its horizontal average. On a horizontally periodic domain, all but two terms vanish in the averaged equation:

$$\partial_t M(x, y, 0, t) = \overline{\mathcal{F}}(0, t). \tag{2.22}$$

Because the forcing is horizontally uniform in all of our simulations, this reduces to (2.13). Note that subtracting (2.22) from (2.21) yields a condition on the integral of A:

$$\frac{if}{2}\nabla^2 \int_{-H}^{0} A\, dz = \mathcal{F}' - \frac{i\zeta}{2} M(x, y, 0, t), \tag{2.23}$$

where $\mathcal{F}' = \mathcal{F} - \overline{\mathcal{F}}$. Note that unlike in YBJ where the integral in left-hand side of Equation 2.23 is set to zero which eliminates the barotropic mode, our boundary conditions allows for a barotropic mode.

2.10 Appendix C: YBJ Potential Energy Budget

The YBJ potential energy budget can be formed by multiplying (2.2) by $i\partial_t M^*/2f$ and adding the complex conjugate. However, a more transparent derivation begins with dotting the gradient of (2.16) with $-\nabla A^*/4$ and adding the complex conjugate (setting $C = 0$ as before). The resulting potential energy budget is:

$$\partial_t \mathcal{P} + J(\psi, \mathcal{P}) + \nabla \cdot \mathbf{F}_\mathcal{P} + \partial_z G_\mathcal{P} = d_\mathcal{P} + \gamma^F_\mathcal{P} + \gamma^R_\mathcal{P} + \gamma^A_\mathcal{P}, \tag{2.24}$$

where

$$\mathcal{P} = \frac{N^2}{4f^2}|\nabla M|^2, \tag{2.25a}$$

$$\mathbf{F}_\mathcal{P} = \frac{if}{8}\left[(\nabla^2 A^*)\nabla A - (\nabla^2 A)\nabla A^*\right], \tag{2.25b}$$

$$G_\mathcal{P} = -\frac{1}{4}\left[\nabla A^* \cdot \left(\partial_t \nabla M + J(\psi, \nabla M)\right) + \text{c.c.}\right], \tag{2.25c}$$

$$d_\mathcal{P} = \frac{\nu}{4}\left[\nabla A^* \cdot \nabla^4 \nabla \partial_z M + \nabla A \cdot \nabla^4 \nabla \partial_z M^*\right], \tag{2.25d}$$

$$\gamma^F_\mathcal{P} = -\frac{1}{4}\left[\nabla A^* \cdot \nabla \partial_z \mathcal{F} + \nabla A \cdot \nabla \partial_z \mathcal{F}^*\right], \tag{2.25e}$$

$$\gamma^R_\mathcal{P} = \frac{i}{8}\left[\nabla A^* \cdot \nabla(\zeta \partial_z M) - \nabla A \cdot \nabla(\zeta \partial_z M^*)\right], \tag{2.25f}$$

$$\gamma^A_\mathcal{P} = \frac{1}{4}\left[\nabla A^* \cdot J(\nabla\psi, \partial_z M) + \nabla A \cdot J(\nabla\psi, \partial_z M^*)\right]. \tag{2.25g}$$

By making a plane wave ansatz in the vertical and using integration by parts, this energy budget can be brought into the same form as the potential energy budget in Rocha et al. (2018). In this paper we are primarily concerned with the domain-integrated potential energy budget. Under domain integration terms which differ by a divergence are the same and hence the interpretation of terms also remains the same. The domain-integrated potential energy budget is (2.6), where again $\mathcal{D}_\mathcal{P} = \langle d_\mathcal{P} \rangle$ and $\Gamma^i_\mathcal{P} = \langle \gamma^i_\mathcal{P} \rangle$. In the case of horizontally uniform forcing, there is no generation of potential energy by the winds and hence $\gamma^F_\mathcal{P} = 0$.

Chapter 3

REGIMES OF NEAR-INERTIAL WAVE DYNAMICS

This chapter is reproduced from the published article:

Conn, S., Callies, J., and Lawrence, A., 2025: Regimes of Near-Inertial Wave Dynamics. *Journal of Fluid Mechanics*, **1002**, A22, doi:10.1017/jfm.2024.1175.

As a self-contained work, some notation may differ from conventions used elsewhere in this thesis.

3.1 Abstract

When atmospheric storms pass over the ocean, they resonantly force near-inertial waves (NIWs), internal waves with a frequency close to the local Coriolis frequency f. It has long been recognised that the evolution of NIWs is modulated by the ocean's mesoscale eddy field. This can result in NIWs being concentrated into anticyclones and provide an efficient pathway for their propagation to depth. We here analyse the eigenmodes of NIWs in the presence of mesoscale eddies and heavily draw on parallels with quantum mechanics. Whether the eddies are effective at modulating the behaviour of NIWs depends on the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$, where λ is the deformation radius and Ψ is a scaling for the eddy streamfunction. If $\varepsilon \gg 1$, NIWs are strongly dispersive, and the waves are only weakly affected by the eddies. We calculate the perturbations away from a uniform wave field and the frequency shift away from f. If $\varepsilon \ll 1$, NIWs are weakly dispersive, and the wave evolution is strongly modulated by the eddy field. In this weakly dispersive limit, the WKB approximation, from which ray tracing emerges, is a valid description of the NIW evolution even if the large-scale atmospheric forcing apparently violates the requisite assumption of a scale separation between the waves and the eddies. The large-scale forcing excites many wave modes, each of which varies on a short spatial scale and is amenable to asymptotic analysis analogous to the semi-classical analysis of quantum systems. The strong modulation of weakly dispersive NIWs by eddies has the potential to modulate the energy input into NIWs from the wind, but we find that this effect should be small under oceanic conditions.

3.2 Introduction

Near-inertial waves (NIWs) play an important role in the global climate system. Being associated with strong vertical shears, they are prone to shear instabilities, which are an important driver of upper ocean mixing (for a review, see Alford et al., 2016). As such, the generation of NIWs is one of the primary mechanisms by which atmospheric storms induce a deepening of the surface mixed layer. This deepening requires mixing with water from below, implicating NIWs in the surface ocean heat budget (Jochum et al., 2013). In the interior of

the ocean, NIWs make up a major fraction of the internal wave kinetic energy (Ferrari and Wunsch, 2009; Alford et al., 2016), and it has been hypothesised that NIW kinetic energy may provide a source of mixing in the deep ocean (Munk and Wunsch, 1998). NIWs might also extract energy from mesoscale eddies (Xie and Vanneste, 2015; Rocha et al., 2018) and hence play a role in the mesoscale energy budget.

In-situ observations of NIWs usually lack significant spatial resolution. The spatial structure of NIWs can generally only be resolved through dedicated field campaigns, for example the Ocean Storms Experiment (D'Asaro, 1985) or the NISKINe field campaign (Voet et al., 2024). Despite this, it has become clear that NIW evolution can be strongly modulated by the presence of mesoscale eddies (e.g., Thomas et al., 2020; Conn et al., 2024). Given the sparsity of NIW observations, theoretical progress has been important in understanding the dynamics of NIWs in the upper ocean.

Early work on NIW–eddy interactions was based on ray tracing theory. Kunze (1985) derived a dispersion relation for NIWs in the presence of a geostrophic background flow. Throughout this paper, we will make the assumption of a barotropic (depth-independent) background flow. The ray tracing equations for a single (flat-bottom) baroclinic mode propagating through such a background flow are

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = \frac{\partial\omega}{\partial\boldsymbol{k}}, \qquad \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}\tau} = -\frac{\partial\omega}{\partial\boldsymbol{x}}, \qquad \omega = \frac{f\lambda^2|\boldsymbol{k}|^2}{2} + \boldsymbol{u}\cdot\boldsymbol{k} + \frac{\zeta}{2}, \tag{3.1}$$

where $\boldsymbol{x} = (x, y)$ is the ray position, τ is time, $\boldsymbol{k}$ is the horizontal wavevector, $\boldsymbol{u}$ is the background velocity, $\zeta = \partial_x v - \partial_y u$ is the background vorticity, and λ is the deformation radius. Here, and throughout the rest of this paper, ω refers to the frequency shift of an NIW away from the local inertial frequency f such that the true frequency is $f + \omega$.

Based on these equations, Kunze (1985) argued that NIWs would be trapped in regions of anticyclonic vorticity where the effective frequency is less than the local f. This trapping arises from the refraction of rays by the background vorticity, i.e., from changes in the wavenumber vector due to spatial gradients of the $\frac{\zeta}{2}$ term in the dispersion relation. Concentration of NIW energy into anticyclones has indeed been observed in the ocean (e.g., Perkins, 1976; Kunze and Sanford, 1984; Thomas et al., 2020; Yu et al., 2022a).

Ray tracing is based on the assumption that the NIWs are propagating through a slowly varying medium. This means that the horizontal scale of the waves has to be much smaller than the scale of the background mesoscale eddy field. Young and Ben-Jelloul (1997, from hereon YBJ) criticised this spatial scale assumption based on the argument that NIWs are forced by large-scale storms and so, at least initially, the waves have a much larger scale than mesoscale eddies. As a remedy, YBJ developed a theory of NIW–eddy interactions that does not rely on the assumption of a spatial scale separation. This was also partly motivated by a desire to explain observations from the Ocean Storms Experiment (D'Asaro et al., 1995), a field campaign that studied the evolution of NIWs in the wake of a large storm in the North

Pacific. A key result of this campaign was that the effect of the mesoscale vorticity on the wave evolution was in clear contradiction with predictions from ray tracing (D'Asaro, 1995).

The YBJ equation describes the evolution of NIWs in the presence of a prescribed geostrophic eddy field while only assuming a temporal scale separation between the inertial period and the characteristic timescale of the eddies. For the barotropic background flow considered throughout this paper, the wave evolution can be split into baroclinic modes that do not interact, so we consider a single baroclinic mode with NIW velocity $[u_w(x, y, t), v_w(x, y, t)]g(z)$, where $g(z)$ is the baroclinic mode structure. The YBJ equation is cast in terms of the variable $\phi = (u_w + iv_w)e^{ift}$, where the factor e^{ift} removes oscillations at the inertial frequency and leaves ϕ to describe the slow evolution of the envelope that modulates the NIWs. For a single mode propagating through a barotropic background flow, the equation becomes:

$$\frac{\partial \phi}{\partial t} + \mathrm{J}(\psi, \phi) + \frac{i\zeta}{2}\phi - \frac{if\lambda^2}{2}\nabla^2\phi = 0, \tag{3.2}$$

where ψ is the background streamfunction, $\zeta = \nabla^2\psi$ is the background vorticity, and $\mathrm{J}(a, b) = \partial_x a\, \partial_y b - \partial_y a\, \partial_x b$ is the Jacobian operator. The second term describes advection of the NIW field by the background flow. The third term is known as the ζ-refraction term and describes refraction of the NIW field by the background vorticity. This term is necessary to obtain concentration of NIWs into regions of anticyclonic vorticity. The last term is responsible for wave dispersion. Here and throughout this paper, we set the meridional gradient of planetary vorticity $\beta = 0$. The YBJ equation can be modified to include β by replacing $\zeta/2$ with $\zeta/2 + \beta y$ in the refraction term. The β-effect has been proposed to explain the observed equatorward propagation of NIWs in the ocean (Anderson and Gill, 1979; Garrett, 2001; Yu et al., 2022a), and it can dominate the overall NIW evolution in regions with weak mesoscale eddies (e.g., D'Asaro et al., 1995). Because mesoscale vorticity gradients typically dominate over β, however, we here restrict ourselves to $\beta = 0$ for simplicity.

Despite both ray tracing and the YBJ equation being used in the NIW literature, it remains unclear how they relate to each other. Ray tracing has been one of the most widely used tools to interpret observations of NIWs. Ray tracing has had qualitative success in describing observed features of NIW evolution, however, we are not aware of any rigorous comparisons between ray tracing predictions and observations. Ray tracing has revealed aspects of NIW dynamics such as trapping in anticyclones along with an associated propagation to depth (Jaimes and Shay, 2010), stalling in cyclones (Oey et al., 2008), and the interplay between NIWs and turbulent dissipation (Kunze et al., 1995; Essink et al., 2022). Non-standard propagation patterns of NIWs in observations have also been explained using ray tracing (e.g., Byun et al., 2010; Chen et al., 2013). The YBJ equation has been used primarily as a tool in theoretical and numerical studies, although there has been some attempt to make connections with observations. Asselin and Young (2020) used simulations of the YBJ equation coupled to a quasi-geostrophic mesoscale eddy field to investigate the sequence of events

that lead to the downward propagation of wind-forced NIWs. Thomas et al. (2020) calculated the NIW wavevector using an expression based on the YBJ equation. The predictions from YBJ were broadly in agreement with observations. Conn et al. (2024) directly used the YBJ equation to interpret NIW observations on a mooring array, showing that it successfully captured the amplitude and phase evolution, including differences across the mooring array caused by mesoscale vorticity gradients. Any comparison of the results of these disparate studies is complicated by the different methods used. A better understanding of the relationship between ray tracing and YBJ would clarify the physical similarities and differences.

Further complicating the picture, observations reveal a varied picture of the importance of the mesoscale vorticity on NIW evolution. During the Ocean Storms Experiment, mesoscale eddies had a muted impact on the NIW field (D'Asaro, 1995), whereas other observational studies found a strong imprint of mesoscale eddies onto the NIW field. For example, Thomas et al. (2020) demonstrated that the evolution of the NIW wavevector was driven by gradients in the mesoscale vorticity during the NISKINe experiment in the North Atlantic. Extending the original argument by YBJ, Thomas et al. (2024a) argued that these differences in the impact of mesoscale vorticity could be explained primarily by differences in the strength of wave dispersion. The stronger dispersion in the Ocean Storms Experiment, they argued, was the result of the forcing projecting onto lower baroclinic modes, a stronger stratification, and weaker eddies. As a result, the effect of refraction by mesoscale vorticity was suppressed in the Ocean Storms Experiment, whereas it was more pronounced in NISKINe.

In this paper, we aim to clarify how ray tracing relates to YBJ dynamics. Given the widespread use of ray tracing in the literature, we aim to understand the conditions under which results from ray tracing are accurate. To this end, we consider the YBJ equation in both a strong and weak-dispersion regime. We begin by providing a simplified treatment of the strong-dispersion regime. Next, we show that the ray tracing equations emerge asymptotically from the YBJ equation in the limit of weak dispersion. Our analysis shows that the WKB approximation and thus ray tracing can be valid even in the presence of a large-scale forcing, despite the YBJ critique. The forcing decomposes into several modes that themselves exhibit small-scale structure. We find the existence of isotropic and anisotropic modes. The isotropic modes are characterised by fast variations along streamlines, while the anisotropic modes have weak variations along streamlines. We discuss the physical processes important in both classes. Finally, we consider how these regimes might modulate the energy injection into the NIW band by the winds, finding that such a modulation is likely weak under oceanic conditions.

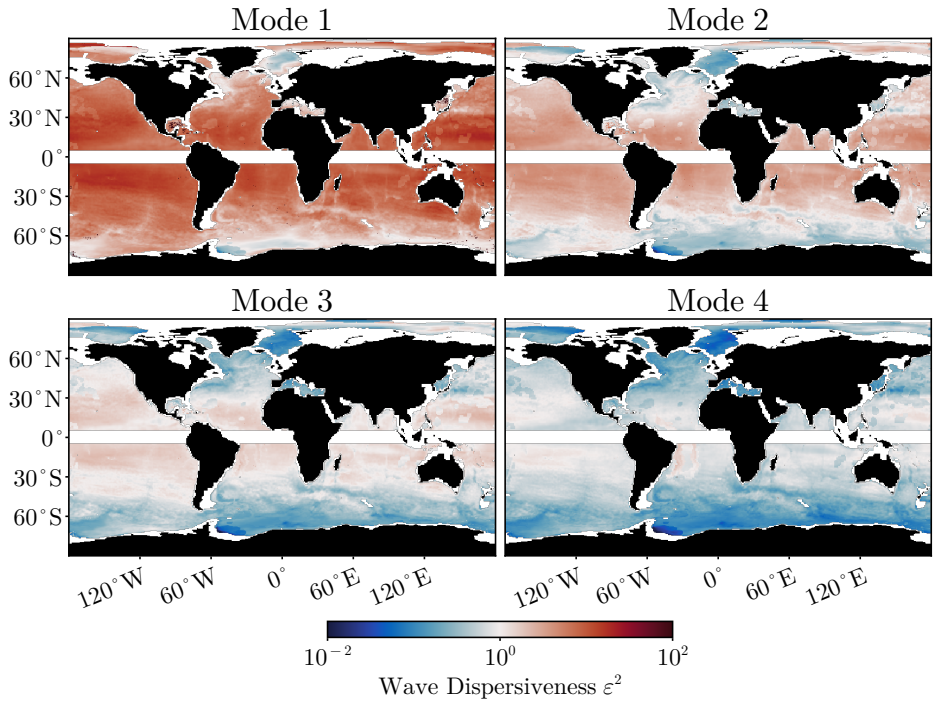

Figure 3.1: Wave dispersiveness $\varepsilon^2 = |f|\lambda^2/\Psi$ plotted throughout the ocean for the first four baroclinic modes, with the deformation radius λ estimated from hydrography and the streamfunction magnitude Ψ from altimetry. The equatorial band is blocked out because the mean flow amplitude cannot be estimated with confidence there.

3.3 The YBJ equation

3.3.1 Decomposition into horizontal modes

We begin by non-dimensionalising the YBJ equation. Given the scalings $x, y \sim L$, $\psi \sim \Psi$ and $t \sim L^2/\Psi$, we obtain the following non-dimensional form of the YBJ equation:

$$\frac{\partial \phi}{\partial t} + \mathrm{J}(\psi, \phi) + \frac{i\zeta}{2}\phi - \frac{i\varepsilon^2}{2}\nabla^2\phi = 0, \tag{3.3}$$

where $\varepsilon^2 = f\lambda^2/\Psi$ is the wave dispersiveness (assuming $f > 0$). For readers familiar with Young and Ben-Jelloul (1997), our ε^2 is equivalent to their Υ^{-1}. We remind the reader that we have assumed a single baroclinic mode, but ε does vary among baroclinic modes through λ. The wave dispersiveness also varies spatially throughout the ocean (Fig. 3.1). We calculate ε for the first four baroclinic modes from observations as described in Appendix 3.10. Except for the high latitudes, the first and second baroclinic modes are almost entirely in the strongly dispersive regime ($\varepsilon \gg 1$). Higher baroclinic modes are to be more weakly dispersive, with $\varepsilon < 1$ almost everywhere for mode 4. For a given baroclinic mode, low-latitude regions are more strongly dispersive, while higher latitudes and western boundary currents are more weakly dispersive.

Note that (3.3) is a Schrödinger equation. This parallel is made clear if we write (3.3) as

$$i\frac{\partial \phi}{\partial t} = H\phi, \qquad H = -\frac{\varepsilon^2}{2}\nabla^2 - i\,\mathrm{J}(\psi, \cdot) + \frac{\zeta}{2}. \tag{3.4}$$

The operator H is known as the Hamiltonian operator. While the presence of first derivatives in the Hamiltonian stemming from advection may be unfamiliar to some, such terms arise in quantum mechanics when describing a charged particle in a magnetic field. This analogy to quantum mechanics was pointed out by Balmforth et al. (1998), and we will here exploit it extensively. Rocha et al. (2018) also used this analogy to derive the equivalent of Ehrenfest's theorem for NIWs, while Danioux et al. (2015) explained the concentration of NIWs into anticyclones via the analogue of quantum conservation laws. While we are setting $\beta = 0$ in this paper, we note that the quantum analogue to $\beta \neq 0$ is known as the "Wannier–Stark ladder," where the potential due to the mesoscale vorticity modulates a linear ramp due to β (Balmforth and Young, 1999).

The operator H is Hermitian, i.e.,

$$\int \varphi^* H\phi \,\mathrm{d}^2\boldsymbol{x} = \int (H\varphi)^* \phi \,\mathrm{d}^2\boldsymbol{x} \tag{3.5}$$

for sufficiently regular functions φ and ϕ so it has real eigenvalues. We also assume H is compact so that the eigenmodes form a complete orthonormal basis. Let $\boldsymbol{\mu}$ label the eigenmodes $\hat{\phi}_{\boldsymbol{\mu}}(x, y)$ and associated eigenvalues $\omega_{\boldsymbol{\mu}}$ of the operator H,

$$H\hat{\phi}_{\boldsymbol{\mu}} = \omega_{\boldsymbol{\mu}}\hat{\phi}_{\boldsymbol{\mu}}. \tag{3.6}$$

The field ϕ can then be expanded in the eigenmode basis as

$$\phi(x, y, t) = \sum_{\boldsymbol{\mu}} a_{\boldsymbol{\mu}}(t)\hat{\phi}_{\boldsymbol{\mu}}(x, y), \tag{3.7}$$

where $a_{\boldsymbol{\mu}}(t)$ is the projection of ϕ onto the eigenmode $\hat{\phi}_{\boldsymbol{\mu}}$. The coefficients $a_{\boldsymbol{\mu}}(t)$ then evolve according to

$$\frac{\mathrm{d}a_{\boldsymbol{\mu}}}{\mathrm{d}t} = -i\omega_{\boldsymbol{\mu}}a_{\boldsymbol{\mu}}, \qquad \text{so} \qquad a_{\boldsymbol{\mu}}(t) = a_{\boldsymbol{\mu}}(0)e^{-i\omega_{\boldsymbol{\mu}}t}. \tag{3.8}$$

Therefore, the eigenvalue $\omega_{\boldsymbol{\mu}}$ represents the frequency shift of the mode away from f. The total dimensional NIW frequency is hence given by $f(1 + \mathrm{Ro}\,\omega_{\boldsymbol{\mu}})$, where $\mathrm{Ro} = \Psi/fL^2$ is the Rossby number. Furthermore, because the eigenvalues are real and the modes are orthogonal, the kinetic energy of the waves is conserved.

We will consider this problem on a doubly periodic domain with size $2\pi \times 2\pi$. This is intended to represent a local view of an ocean that is filled with a random sea of eddies. The solutions we calculate are perfectly periodic and extend across all eddies. In reality, of course, the background field is not perfectly periodic and this causes the solutions to become localised in certain regions. Therefore, the solutions we calculate on the $2\pi \times 2\pi$ should be thought of similarly.

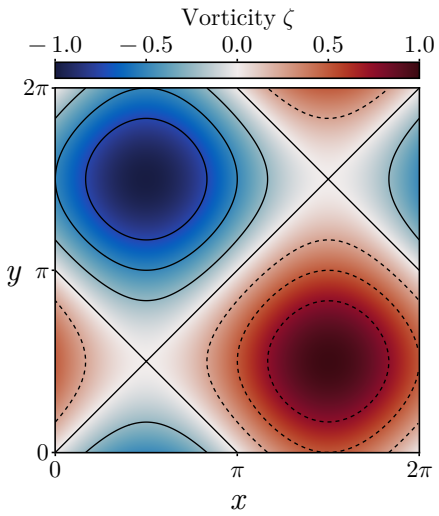

Figure 3.2: Dipole vorticity with an anticyclone in the upper left corner and a cyclone in the lower right corner. The contours depict the streamfunction with positive values denoted by solid lines and negative values denoted by dashed lines.

As a key example in this paper, we will consider a $2\pi \times 2\pi$ domain that contains a dipole vortex given by (figure 3.2; *cf.,* Asselin et al., 2020)

$$\psi = \frac{1}{2}\left(\sin x - \sin y\right). \tag{3.9}$$

The analysis below is general, however, and can be applied to more general background flows.

3.3.2 Numerical calculation of eigenvalues and eigenmodes

For most choices of the background flow ψ, analytical solutions for the eigenfunctions of H do not exist and numerical solutions are required. Solving the eigenvalue equation numerically requires us to discretise the operator H. The discrete eigenfunction is expressed as a vector, and the problem reduces to finding the eigenvalues of a finite matrix. The operator H is Hermitian, and so it is desirable for any discrete representation of H to also be Hermitian. A fourth-order central finite difference scheme for the Laplacian term preserves this property. More care is required for the advection operator, for which we use the enstrophy-conserving scheme from Arakawa (1966) to preserve the Hermitian nature of the operator and guarantee that the eigenvalues of the matrix are real. Having real eigenvalues ensures that the conservation of NIW kinetic energy is respected in the discrete system. The exact method of numerically solving the eigenvalue problem is detailed in Appendix 3.11.

3.4 The strong-dispersion limit

The limit $\epsilon \gg 1$ is known as the strong-dispersion limit. YBJ showed that in this limit, the solution to the YBJ equation becomes proportional to the streamfunction ψ. They additionally showed that frequency shifts away from f are proportional to the domain-averaged

kinetic energy of the mesoscale flow. These same results can be derived by considering the eigenvalue problem posed above; see (3.17) and (3.20) for the result. In our framework, we can additionally derive information about the next-order perturbations to the NIW field; see (3.24) and (3.25) below.

When ε is large, we split the operator H into two parts

$$H = \varepsilon^2 H^{(0)} + H^{(1)}, \tag{3.10}$$

where $H^{(0)} = -\frac{1}{2}\nabla^2$ and $H^{(1)} = \frac{1}{2}\zeta - i\,\mathrm{J}(\psi, \cdot)$. Because $\varepsilon^2 \gg 1$, this implies $H^{(1)}$ is a small correction to $\varepsilon^2 H^{(0)}$, and perturbation theory can be used to solve this system. We expand both $\hat{\phi}_{\boldsymbol{\mu}}$ and $\omega_{\boldsymbol{\mu}}$ in powers of ε^{-2}:

$$\hat{\phi}_{\boldsymbol{\mu}} = \sum_{n=0}^{\infty} \varepsilon^{-2n} \hat{\phi}_{\boldsymbol{\mu}}^{(n)}, \qquad \omega_{\boldsymbol{\mu}} = \varepsilon^2 \sum_{n=0}^{\infty} \varepsilon^{-2n} \omega_{\boldsymbol{\mu}}^{(n)}. \tag{3.11}$$

At $O(\varepsilon^2)$ the eigenvalue problem is

$$H^{(0)} \hat{\phi}_{\boldsymbol{\mu}}^{(0)} = \omega_{\boldsymbol{\mu}}^{(0)} \hat{\phi}_{\boldsymbol{\mu}}^{(0)}, \tag{3.12}$$

where $\hat{\phi}_{\boldsymbol{\mu}}^{(0)}$ is the eigenfunction of the unperturbed problem with eigenvalue $\omega_{\boldsymbol{\mu}}^{(0)}$. We assume the domain is doubly periodic and goes from 0 to 2π in x and y. The solution is

$$\hat{\phi}_{\boldsymbol{\mu}}^{(0)} = e^{i\boldsymbol{\mu}\cdot\boldsymbol{x}}, \qquad \omega_{\boldsymbol{\mu}}^{(0)} = \frac{|\boldsymbol{\mu}|^2}{2}, \tag{3.13}$$

where $\boldsymbol{\mu}$ is a two-dimensional vector with integer components, such that the eigenfunctions are plane waves in x and y.

3.4.1 The leading-order mode

NIWs are forced by atmospheric storms, which have a much larger horizontal scale than mesoscale eddies and can be idealised as a uniform forcing. We assume that the result of this forcing is the excitation of a constant non-zero ϕ. The projection of this initial condition onto a given mode can thus be found by integrating that mode across the domain. For plane waves, a domain integral will vanish unless $\boldsymbol{\mu} = 0$, such that a uniform forcing will only project onto the $\boldsymbol{\mu} = 0$ mode in the unperturbed case. We begin by focusing on that case to obtain expressions for the perturbations to its spatial structure as well as its frequency shift. A small part of the forcing, however, projects onto modes with $\boldsymbol{\mu} \neq 0$, and we will return to these higher modes below.

The leading-order solution for $\boldsymbol{\mu} = 0$ is $\hat{\phi}_0^{(0)} = 1$ and $\omega_0^{(0)} = 0$, and there is no modulation of the waves by the mesoscale eddy field. To obtain this modulation, we must go to higher order. At $O(\varepsilon^0)$, the eigenvalue problem is

$$H^{(0)} \hat{\phi}_0^{(1)} + H^{(1)} \hat{\phi}_0^{(0)} = \omega_0^{(0)} \hat{\phi}_0^{(1)} + \omega_0^{(1)} \hat{\phi}_0^{(0)}. \tag{3.14}$$

With $\omega_0^{(0)} = 0$ and the advection term in $H^{(1)}$ vanishing when acting on $\hat{\phi}_0^{(0)} = 1$, this reduces to

$$-\frac{1}{2}\nabla^2\hat{\phi}_0^{(1)} + \frac{\zeta}{2} = \omega_0^{(1)}. \tag{3.15}$$

The two terms on the left vanish when integrated over the doubly periodic domain, so we conclude that $\omega_0^{(1)} = 0$.

There is, however, a correction to the eigenfunction at this order, determined by

$$\nabla^2\hat{\phi}_0^{(1)} = \nabla^2\psi. \tag{3.16}$$

With periodic boundary conditions, the solution to this is

$$\hat{\phi}_0^{(1)} = \psi, \tag{3.17}$$

where we have assumed that ψ is defined such that it has zero domain average. This recovers the expression for $\hat{\phi}$ from YBJ. The structure of the mesoscale eddy field is imprinted onto the waves by the $\varepsilon^{-2}\hat{\phi}_0^{(1)}$ term. Because the modulation is by the real streamfunction ψ, only the NIW amplitude is modulated by mesoscale eddies. The NIW field remains in phase across the domain.

We now also seek the leading non-zero correction to the eigenvalue, for which we go up another order. The eigenvalue equation at $O(\varepsilon^{-2})$ is

$$H^{(0)}\hat{\phi}_0^{(2)} + H^{(1)}\hat{\phi}_0^{(1)} = \omega_0^{(0)}\hat{\phi}_0^{(2)} + \omega_0^{(1)}\hat{\phi}_0^{(1)} + \omega_0^{(2)}\hat{\phi}_0^{(0)}. \tag{3.18}$$

With $\omega_0^{(0)} = \omega_0^{(1)} = 0$ and $\mathrm{J}(\psi, \psi) = 0$, this simplifies to

$$-\frac{1}{2}\nabla^2\hat{\phi}_0^{(2)} + \frac{1}{2}\psi\nabla^2\psi = \omega_0^{(2)}. \tag{3.19}$$

The first term on the left vanishes under domain integration. Integrating the second term on the left by parts yields

$$\omega_0^{(2)} = -\frac{1}{2}\frac{\int |\nabla\psi|^2 \,\mathrm{d}^2\boldsymbol{x}}{\int \mathrm{d}^2\boldsymbol{x}}. \tag{3.20}$$

The leading-order frequency shift is $\varepsilon^{-2}\omega_0^{(2)}$. Given that $\varepsilon^{-2} \ll 1$, the frequency shift away from f is suppressed substantially, even compared to the small frequency shift assumed from the outset. Re-dimensionalising the expression results in

$$\omega_0^{(2)} = -\frac{1}{2f_0\lambda^2}\frac{\int |\nabla\psi|^2 \,\mathrm{d}^2\boldsymbol{x}}{\int \mathrm{d}^2\boldsymbol{x}}. \tag{3.21}$$

This agrees with the YBJ result for the dispersion relation in the strong-dispersion regime, indicating that the frequency shift is proportional to the average kinetic energy of the eddy field.

3.4.2 Higher-order modes

We now return to the higher modes with $\boldsymbol{\mu} \neq 0$. These modes are degenerate to leading order. For example, the modes $(1, 0)$, $(-1, 0)$, $(0, 1)$ and $(0, -1)$ all have $\omega_{\boldsymbol{\mu}}^{(0)} = \frac{1}{2}$. Degenerate perturbation theory is necessary to calculate the first-order corrections to the eigenvalues and eigenfunctions (e.g., Sakurai and Napolitano, 2020). To obtain these corrections, we will proceed naively with the calculation. We will run into a contradiction that motivates us to choose a different basis set than was chosen in (3.13). To those familiar with degenerate perturbation theory, this may seem unnecessary, but we believe it to be more pedagogical.

We again start from the $O(\varepsilon^0)$ equation, which now reads

$$H_0\hat{\phi}_{\boldsymbol{\mu}}^{(1)} + H_1\hat{\phi}_{\boldsymbol{\mu}}^{(0)} = \omega_{\boldsymbol{\mu}}^{(0)}\hat{\phi}_{\boldsymbol{\mu}}^{(1)} + \omega_{\boldsymbol{\mu}}^{(1)}\hat{\phi}_{\boldsymbol{\mu}}^{(0)}. \tag{3.22}$$

Multiplying this equation by $\hat{\phi}_{\boldsymbol{\nu}}^{(0)*}$, with both $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ labelling one of the modes in the degenerate group, and integrating over the domain results in

$$\int \hat{\phi}_{\boldsymbol{\nu}}^{(0)*}\left(H_0 - \omega_{\boldsymbol{\mu}}^{(0)}\right)\hat{\phi}_{\boldsymbol{\mu}}^{(1)}\,\mathrm{d}^2\boldsymbol{x} = \omega_{\boldsymbol{\mu}}^{(1)}\int \hat{\phi}_{\boldsymbol{\nu}}^{(0)*}\hat{\phi}_{\boldsymbol{\mu}}^{(0)}\,\mathrm{d}^2\boldsymbol{x} - \int \hat{\phi}_{\boldsymbol{\nu}}^{(0)*}H_1\hat{\phi}_{\boldsymbol{\mu}}^{(0)}\,\mathrm{d}^2\boldsymbol{x}. \tag{3.23}$$

Using integration by parts, the H_0 on the left can be swapped for $\omega_{\boldsymbol{\nu}}^{(0)}$. Because the modes are degenerate to this order, the left-hand side vanishes. Furthermore, using the orthonormality of the eigenfunctions, the corrections to the eigenvalues are determined by

$$\omega_{\boldsymbol{\mu}}^{(1)}\delta_{\boldsymbol{\nu}\boldsymbol{\mu}} = \frac{1}{4\pi^2}\int \hat{\phi}_{\boldsymbol{\nu}}^{(0)*}H_1\hat{\phi}_{\boldsymbol{\mu}}^{(0)}\,\mathrm{d}^2\boldsymbol{x}. \tag{3.24}$$

We have now arrived at our contradiction. The left-hand side of this equation is diagonal, whereas the right-hand side is not necessarily so. In (3.13), we chose a basis for the unperturbed eigenfunctions: $\{e^{ix}, e^{-ix}, e^{iy}, e^{-iy}\}$ for $|\boldsymbol{\mu}| = 1$. The key to degenerate perturbation theory is to choose a basis of the degenerate space to avoid this contradiction. It is clear that the correct basis must diagonalise H_1, which our original choice does not.

To proceed, we calculate the right-hand side of (3.24) in the original basis. This results in a 4×4 matrix. We diagonalise this matrix and find the corresponding linear combination of the original basis functions that diagonalises H_1. The corresponding eigenfunction corrections can be found by solving the screened Poisson equation obtained from the first-order equation (3.22).

$$\left(H_0 - \omega_{\boldsymbol{\mu}}^{(0)}\right)\hat{\phi}_{\boldsymbol{\mu}}^{(1)} = \left(\omega_{\boldsymbol{\mu}}^{(1)} - H_1\right)\hat{\phi}_{\boldsymbol{\mu}}^{(0)}, \tag{3.25}$$

where the $\hat{\phi}_{\boldsymbol{\mu}}^{(0)}$ should be in the basis diagonalising H_1. If H_1 identically vanishes in this subspace, the degeneracy must be lifted at the next order, as in the example below.

3.4.3 Dipole flow solutions

We now consider the specific example of the dipole flow (3.9). Numerical solutions for $\varepsilon = 2$ show that a uniform initial condition projects strongly (98.5% of the energy) onto the $\hat{\phi}_0$ mode (figure 3.3). There is a small but negative frequency shift of $\omega_0 = -0.03104$. This agrees excellently with the predicted frequency shift from (3.20) of $\varepsilon^{-2}\omega_0^{(2)} = \frac{1}{32} = -0.03125$. Additionally, there is weak horizontal structure that aligns with the streamfunction as expected. The root-mean-squared error between the numerical eigenmode and the analytical eigenmode $\hat{\phi}_0^{(0)} + \varepsilon^{-2}\hat{\phi}_0^{(1)} = 1 + \varepsilon^{-2}\psi$ is 1%. The agreement is excellent despite ε not being particularly large.

For the dipole flow, the right-hand side of (3.24) is zero for all combinations of basis functions of the $\omega_{\boldsymbol{\mu}}^{(0)} = \frac{1}{2}$ subspace. Therefore, there are no first-order frequency shifts, $\omega_{\boldsymbol{\mu}}^{(1)} = 0$, and the degeneracy is not lifted at this order. Performing the same procedure that led to (3.24) on the second-order equation yields

$$\omega_{\boldsymbol{\mu}}^{(2)}\delta_{\boldsymbol{\nu}\boldsymbol{\mu}} = \frac{1}{4\pi^2}\int \hat{\phi}_{\boldsymbol{\nu}}^{(0)*} H_1 \hat{\phi}_{\boldsymbol{\mu}}^{(1)}\,\mathrm{d}^2\boldsymbol{x}. \tag{3.26}$$

For our trial basis consisting of the four plane waves, we solve the screened Poisson equation (3.25) for the corresponding $\hat{\phi}_{\boldsymbol{\mu}}^{(1)}$. This is tedious but doable because the right-hand side is just a sum of sines and cosines. The equation for the second-order frequency shift can be diagonalised, and this time the eigenvalues are not zero and the degeneracy is lifted. We find for $\omega_{\boldsymbol{\mu}}^{(2)}$ the values $-\frac{1}{96}$, $-\frac{7}{96}$, $-\frac{49}{96}$, and $-\frac{55}{96}$, only the first of which corresponds to an eigenfunction that the forcing projects onto at this order. The leading-order eigenfunction of that mode is $\hat{\phi}_{\boldsymbol{\mu}}^{(0)} = -\psi$ (figure 3.3). The eigenvalue $\varepsilon^2\omega_{\mu}^{(0)} + \varepsilon^{-2}\omega_{\boldsymbol{\mu}}^{(2)} = 1.99739$ is again in excellent agreement with the numerical eigenvalue of 1.99729.

In this regime, horizontal structure in the waves primarily arises due to $\hat{\phi}_0^{(1)}$, which is suppressed by $O(\varepsilon^{-2})$. There is also horizontal structure due to modes with $\boldsymbol{\mu} \neq 0$, but these are projected onto weakly; the fraction of the variance accounted for by such a mode is $O(\varepsilon^{-4})$ (Sakurai and Napolitano, 2020). As such, the wave potential energy, which depends on horizontal gradients in the wave field, is also suppressed. Xie and Vanneste (2015) associated the generation of wave potential energy with a sink of the background eddy kinetic energy in a process known as stimulated generation. Given the weak generation of horizontal structure, stimulated generation is weak in the strong-dispersion regime.

3.5 The weak-dispersion limit

The limit $\varepsilon \ll 1$ is known as the weak-dispersion limit. Because ε^2 multiplies the highest-order derivative in the eigenvalue equation, the limit $\varepsilon \to 0$ is a singular perturbation problem. Before addressing the general problem, we build intuition with two simple examples. These examples suggest that there are two classes of modes. One class is characterised by waves that vary slowly along the streamlines of the background flow and more rapidly across streamlines; they are captured by an anisotropic scaling of the wavenumber with ε. The other

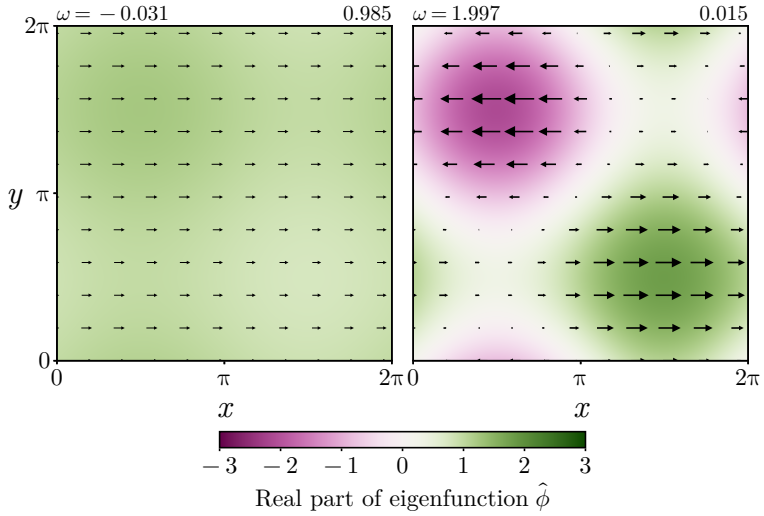

Figure 3.3: Numerical solution to the eigenvalue problem (3.6) with $\varepsilon = 2$ for the dipole flow. A uniform forcing primarily projects onto the mode shown on the left, with the mode with the second-highest projection shown on the right only making up less than 2% of the energy. The eigenvalues ω and projection fractions (of energy) are shown in the panel titles. Vectors show the corresponding NIW velocities.

class has even faster variations in both directions and requires an isotropic scaling. We develop a uniformly valid approximation that captures both of these classes.

3.5.1 Parallel shear flow

We begin with an example of a parallel shear flow in which the streamfunction ψ is a function of x only. The symmetry in y means the problem reduces to a one-dimensional eigenvalue problem. Balmforth et al. (1998) considered this problem for a specific example of a shear flow that can be solved in closed form. Zhang and Xie (2023) considered the limits of strong and weak dispersion for the same mean flow. Here, we address how the weak-dispersion limit can be analysed for a general parallel shear flow and apply the procedure to the example flow from Balmforth et al. (1998). Our goal is to calculate the structure of the eigenmodes and their corresponding eigenvalues. We begin by introducing the WKB method from which the two scalings arise. For the anisotropic scaling, the key result is (3.34); for the isotropic scaling, the equivalent result is (3.38).

We assume that the streamfunction $\psi(x)$ is periodic on the domain $[-\pi, \pi]$. The eigenvalue problem (3.6) reduces to

$$-\frac{\varepsilon^2}{2}\nabla^2\hat{\phi} - iv\frac{\partial\hat{\phi}}{\partial y} + \frac{\zeta}{2}\hat{\phi} = \omega\hat{\phi}, \tag{3.27}$$

where $\zeta = \nabla^2\psi$ and $v = \partial_x\psi$ are both functions of x only, and we have suppressed the label on the eigenmode. The coefficients are independent of y, which motivates the ansatz $\hat{\phi} = \Phi(x)e^{imy}$. Given that the domain has width 2π in y, the wavenumber m must be an

integer. With this ansatz, we are left with the one-dimensional eigenvalue problem

$$-\frac{\varepsilon^2}{2}\frac{\mathrm{d}^2\Phi}{\mathrm{d}x^2}+\left(\frac{\varepsilon^2 m^2}{2}+mv+\frac{\zeta}{2}\right)\Phi=\omega\Phi. \tag{3.28}$$

This is the Schrödinger equation of a particle in one-dimensional potential, with the bracketed term playing the role of the potential (Balmforth et al., 1998).

As ε is small, WKB analysis can be used to find approximations to the eigenvalues and eigenfunctions (e.g., Bender and Orszag, 1999). In WKB theory, the field Φ is expanded as

$$\Phi(x)=\exp\frac{1}{\delta}\sum_{j=0}^{\infty}\delta^j S_j(x), \tag{3.29}$$

where $\delta \ll 1$ is a scaling parameter that we are yet to determine. Substituting this into (3.28) yields

$$-\frac{\varepsilon^2}{2}\left[\frac{1}{\delta^2}\left(\sum_{j=0}^{\infty}\delta^j\frac{\mathrm{d}S_j}{\mathrm{d}x}\right)^2+\frac{1}{\delta}\sum_{j=0}^{\infty}\delta^j\frac{\mathrm{d}^2S_j}{\mathrm{d}x^2}\right]+\frac{\varepsilon^2m^2}{2}+mv+\frac{\zeta}{2}=\omega. \tag{3.30}$$

If we assume $m \sim O(1)$, both the refraction term and the advection terms are $O(1)$, and they must be balanced by a dispersion term of the same order. Requiring the lowest-order dispersion term to be $O(1)$ implies $\delta=\varepsilon$, and the $O(1)$ equation becomes

$$-\frac{1}{2}\left(\frac{\mathrm{d}S_0}{\mathrm{d}x}\right)^2+mv+\frac{\zeta}{2}=\omega. \tag{3.31}$$

By writing $\varepsilon^{-1}\mathrm{d}S_0/\mathrm{d}x = ik$, this equation is analogous to the dispersion relation (3.1) specialised to this parallel shear flow. The function S_0 is found to be

$$S_0(x)=\pm\sqrt{2}i\int^x\sqrt{\omega-mv(x')-\frac{\zeta(x')}{2}}\,\mathrm{d}x' \tag{3.32}$$

and determines the leading-order phase variations of the solution. One can additionally show (Bender and Orszag, 1999, equation 10.1.12) that the next-order solution is

$$S_1(x)=-\frac{1}{4}\ln\left(\omega-mv-\frac{\zeta}{2}\right), \tag{3.33}$$

which determines the leading-order amplitude modulation of the solution.

This asymptotic expansion is valid away from regions where the integrand above is zero. These are known as turning points of the problem and exist if $\omega < \max(mv+\zeta/2)$. The associated eigenfunctions are referred to as bound states. Near turning points, $\omega-mv-\zeta/2$ can be approximated by a linear function of x, and solutions to (3.31) are given by Airy functions. The Airy function solutions must be asymptotically matched to the solutions away from the turning points. This yields an integral constraint from which the eigenvalues ω can be determined. The problem as formulated above is the classic two-turning point problem,

and the asymptotic matching procedure is well documented (e.g., Bender and Orszag, 1999, equation 10.5.6). The resulting condition for ω, often referred to as a quantisation condition, is

$$\frac{\sqrt{2}}{\varepsilon}\int_{x_0}^{x_1}\sqrt{\omega - mv(x) - \frac{\zeta(x)}{2}}\,\mathrm{d}x = \left(n+\frac{1}{2}\right)\pi, \quad \text{with} \quad n = 0, 1, \dots, \tag{3.34}$$

where x_0 and x_1 are the turning points. The projection of a uniform forcing onto these modes can also be calculated asymptotically. The domain integral of a mode is dominated by contributions from the turning points (e.g., Bender and Orszag, 1999, equation 10.4.24).

If $\omega > \max(\zeta/2 + mv)$ then there are no turning points. The corresponding eigenmodes are referred to as free states, and the quantisation condition is replaced by

$$\frac{\sqrt{2}}{\varepsilon}\int_{-\pi}^{\pi}\sqrt{\omega - mv(x) - \frac{\zeta(x)}{2}}\,\mathrm{d}x = 2n\pi, \quad \text{with} \quad n = 0, 1, \dots \tag{3.35}$$

Note the lack of a half-integer shift that for bound states arises from the Airy behaviour near turning points. The lack of turning points in the free states also means (3.32) is valid across the entire domain. Because the eigenfunctions of these free states are oscillatory in the entire domain, a uniform forcing projects only weakly onto them, and we do not discuss them any further. We also note that the discretisation of the free states is due to the periodic domain; they would be replaced by a continuum of modes in an infinite domain, while the bound states would remain discrete.

Under this scaling, the WKB modes are anisotropic. We assumed $m \sim O(1)$, which means that the modes' phase varies in y on a length scale $O(1)$. In contrast, the leading-order phase variations in x come from $\varepsilon^{-1}S_0$ and therefore occur on a scale $O(\varepsilon)$. The phase varies slowly along streamlines and rapidly across streamlines. This makes refraction and advection come in at the same order as cross-streamline dispersion.

An alternative would be to choose the scaling $m \sim O(\varepsilon^{-2})$. Repeating the WKB ansatz requires a choice of $\delta = \varepsilon^2$ and $\omega \sim O(\varepsilon^{-2})$ in order to end up with an equation of a similar form to (3.31):

$$-\frac{1}{2}\left(\frac{\mathrm{d}S_0}{\mathrm{d}x}\right)^2 + \frac{\varepsilon^4 m^2}{2} + \varepsilon^2 mv = \varepsilon^2\omega. \tag{3.36}$$

With the scaling given above, each term is $O(1)$. We can solve for S_0 and the corresponding quantisation condition for bound modes:

$$S_0(x) = \pm\sqrt{2}i\varepsilon\int^{x}\sqrt{\omega - \frac{\varepsilon^2 m^2}{2} - mv(x')}\,\mathrm{d}x', \tag{3.37}$$

$$\frac{\sqrt{2}}{\varepsilon}\int_{x_0}^{x_1}\sqrt{\omega - \frac{\varepsilon^2 m^2}{2} - mv(x)}\,\mathrm{d}x = \left(n+\frac{1}{2}\right)\pi. \tag{3.38}$$

These modes are isotropic. The phase variations in y occur on a scale $O(\varepsilon^2)$, which is the same as in x because phase variations in x now come from $\varepsilon^{-2}S_0$. This makes advection and along-streamline dispersion come in at the same order, and it makes refraction negligible.

Despite the different characteristics of the two scalings, they lead to similar quantisation conditions that differ only by what terms are included. We can combine them into a uniformly valid quantisation condition:

$$\frac{\sqrt{2}}{\varepsilon}\int_{x_0}^{x_1}\sqrt{\omega-\frac{\varepsilon^2m^2}{2}-mv(x)-\frac{\zeta(x)}{2}}\,\mathrm{d}x=\left(n+\frac{1}{2}\right)\pi. \tag{3.39}$$

The "potential" governing the wave evolution is therefore

$$V(x)=\frac{\varepsilon^2m^2}{2}+mv(x)+\frac{\zeta(x)}{2}. \tag{3.40}$$

Under the anisotropic scaling $m\sim O(1)$, the along-streamline dispersion term is suppressed by a factor ε^2, leaving the $O(1)$ refraction and advection terms to dominate. Under the isotropic scaling $m\sim O(\varepsilon^{-2})$, the advection and along-streamline dispersion terms are enhanced by a factor ε^{-2} and dominate over a now negligible refraction term. In both cases, the general equation is obtained by retaining a term that is of higher order, which is allowed in an asymptotic theory. A uniform forcing only projects onto modes with $m=0$, so all the modes projected onto are of the anisotropic variety.

We now consider a specific example of a parallel shear flow that varies sinusoidally in x:

$$\psi=\cos x. \tag{3.41}$$

This shear flow has a region of anticyclonic vorticity at the centre of the domain and cyclonic vorticity centred on $x=\pm\pi$ (figure 3.4a,b). This is a rare example in which the eigenvalue problem (3.28) can be solved exactly using Mathieu functions (Balmforth et al., 1998). The generally applicable WKB theory described above accurately predicts the eigenvalues, even for a modestly small $\varepsilon=\frac{1}{4}$ (figure 3.4c). We provide the analytical solutions to the WKB integrals in Appendix 3.12. We also note that the symmetry of the problem means that a uniform wind forcing only projects onto modes with even n.

For $m=0$, the eigenmodes are shaped by the potential $V=\frac{\zeta}{2}$ (figure 3.5a). Where $\omega>V$, S_0 is imaginary and the solutions are oscillatory; where $\omega<V$, S_0 is real and the solutions are decaying (figure 3.5a). Near the anticyclonic centre of the flow, the potential is at its lowest and all the modes are oscillatory. Moving further out into the cyclonic region, more and more of the modes become evanescent.

The dependence of the potential on the vorticity ζ leads to trapping of NIW in anticyclones. The trapping arises from the dephasing of the modes that make up the initial condition. This is analogous to the argument in Gill (1984) regarding the vertical propagation of NIWs due to the β-effect. The evolution occurs in three phases (figure 3.5b). First, refraction imprints the mesoscale vorticity onto the initially uniform wave phase, leaving the amplitude unchanged (*cf.,* Asselin et al., 2020). Second, once these phase gradients are sufficiently pronounced, cross-streamline dispersion becomes important and concentrates the wave energy into the

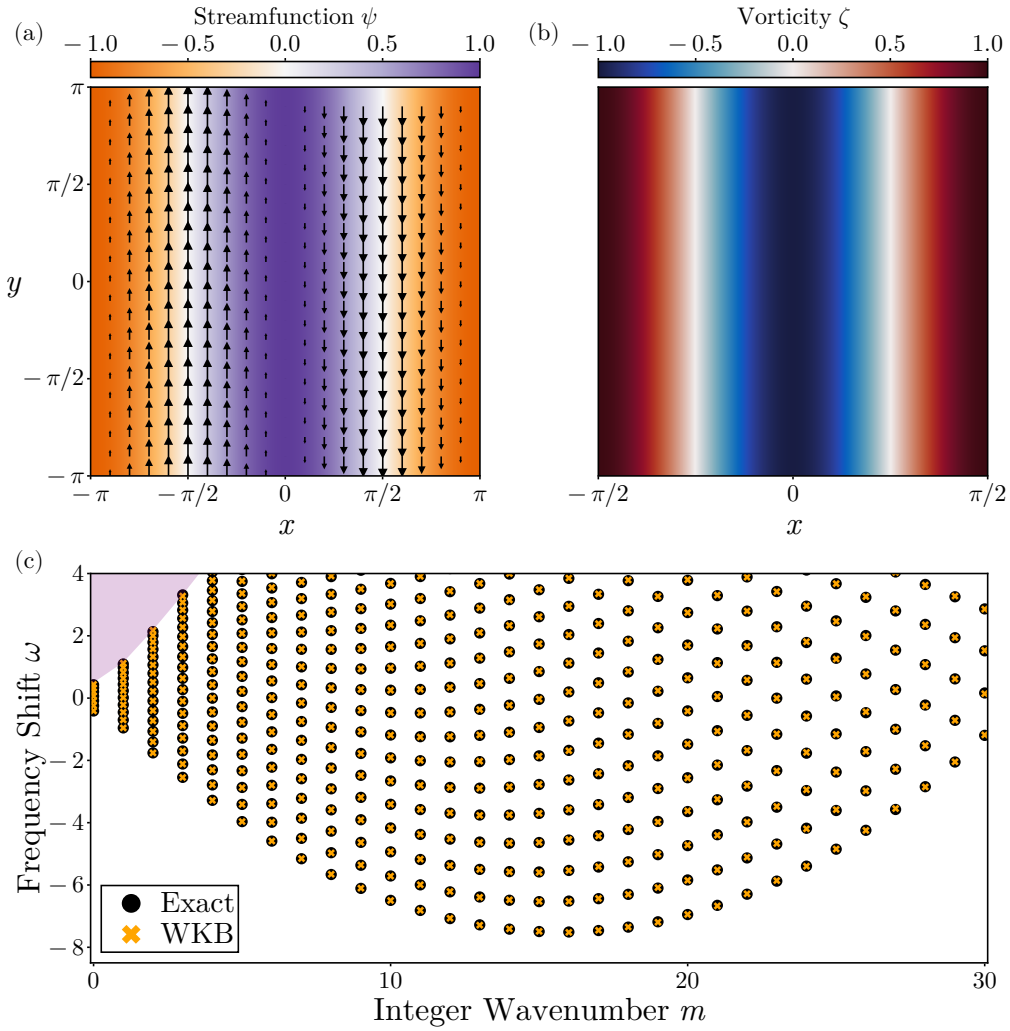

Figure 3.4: (a) Streamfunction and flow vectors for the shear flow example. (b) Vorticity, showing the anticyclonic vorticity concentrated in the centre of the domain and cyclonic vorticity on the outside. (c) Eigenvalues ω as a function of the integer wavenumber m for $\varepsilon = \frac{1}{4}$. The results from WKB theory (orange crosses) are shown along with the exact eigenvalues found from numerical solutions (black circles). The WKB results agree remarkably well with the numerical results, although there are some spurious eigenvalues near the boundary between free and bound modes. The purple shading indicates the region where free modes exist, which are not shown here.

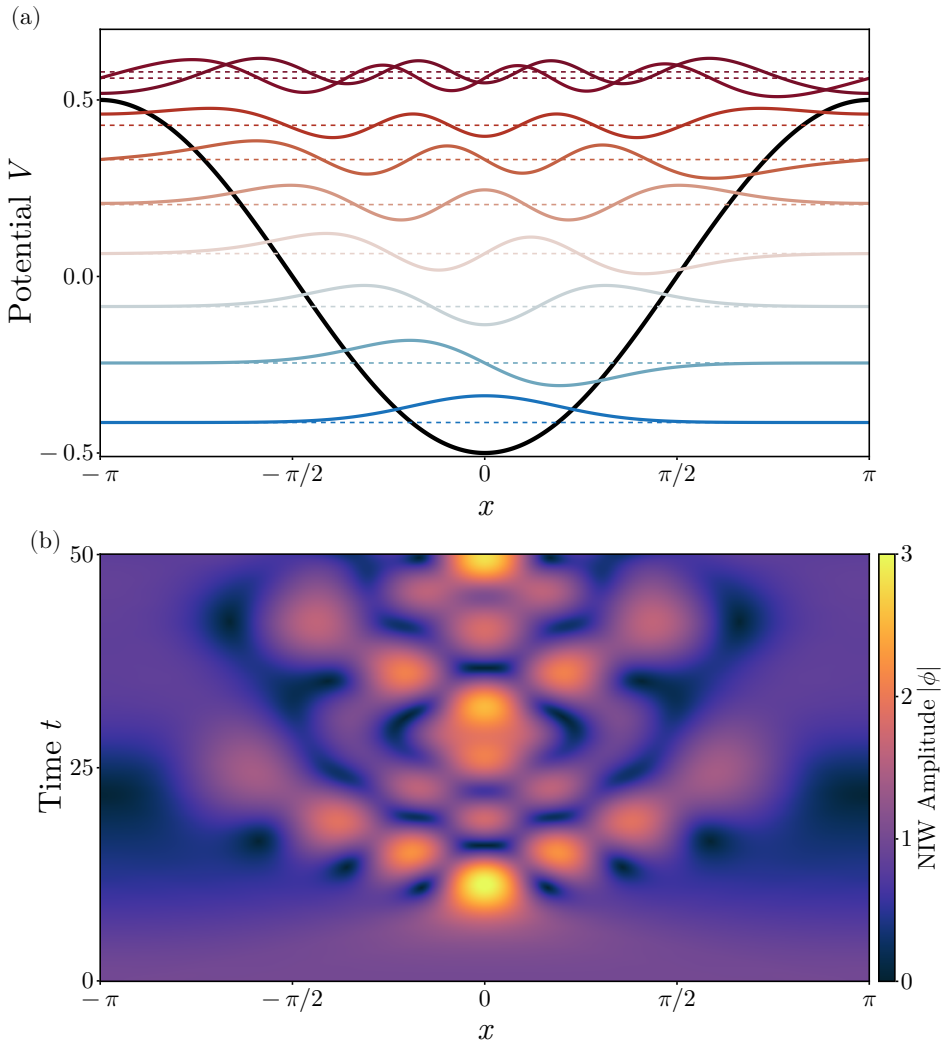

Figure 3.5: (a) The potential $V = \frac{\zeta}{2}$ (black line) of the parallel shear flow for the $m = 0$ mode. The dashed lines show the level of each eigenvalue ω for $\varepsilon = \frac{1}{4}$. The solid coloured lines represent a scaled representation of each eigenfunction corresponding to a given eigenvalue, as identified by the colours. (b) Time evolution of the NIW amplitude $|\phi|$ for the parallel shear flow example with $\varepsilon = \frac{1}{4}$, starting by a uniform field.

centre of the anticyclone, and the wave field assumes a spatial scale $O(\varepsilon)$. Third, the amplitude remains elevated on average within the anticyclonic region but is more spread-out than during the initial concentration. It is this long-time behaviour that corresponds to the fully dephased eigenmodes. The time it takes for this dephasing to occur depends inversely on the spacing of the eigenvalues ω. As ε decreases, the eigenvalues become more finely spaced. The $m = 0$ modes have a spacing $O(\varepsilon)$, so it takes $t \sim O(\varepsilon^{-1})$ for them to dephase. Another way to think of this is that as ε decreases, dispersion becomes weaker, and it takes longer for phase gradients to build up to a level where dispersion is important.

Finally, we note that Asselin et al. (2020) discussed solutions to the YBJ equation for which phase lines are aligned with streamlines and straining is ineffective in driving a decrease in the spatial scale of the wave. That analysis sets the dispersion term to zero, however, and so only captures the initial phase in which refraction dominates. The WKB theory presented above shows that cross-streamline dispersion is of leading order and should not be dropped if the long-term evolution is of interest (*cf.*, figure 3.5b). Our $m = 0$ anisotropic modes can thus be understood as a generalisation of Asselin et al.'s solution. The ineffectiveness of straining due to the alignment of the wave phase with streamlines remains apparent, but cross-streamline dispersion is now taken into account such that the solution remains valid at late times. We further note that our anisotropic modes also allow for slow variations of the wave field along streamlines, such that advection assumes the same importance as refraction and cross-streamline dispersion. These anisotropic modes with $m > 0$ may be excited by a non-uniform forcing, such as a passing atmospheric front (*cf.*, Thomas et al., 2017).

3.5.2 Axisymmetric flow

We now consider a streamfunction with axial symmetry, such that $\psi = \psi(r)$, where r is the radial distance from the origin. Llewellyn Smith (1999) studied NIWs with azimuthal wavenumber zero in an axisymmetric vortex and provided asymptotic expressions for the frequency of the lowest radial mode. Kafiabad et al. (2021) studied a similar case but also considered the impact of NIWs back on the vortex. Using WKB theory, we consider NIWs with an arbitrary azimuthal wavenumber and provide a transcendental equation that can be solved for their frequency as for the parallel shear flows above.

We make the ansatz $\hat{\phi} = A(r)e^{im\theta}$, where θ is the azimuthal angle, and again we drop the mode label. In polar coordinates, (3.6) then reduces to

$$-\frac{\varepsilon^2}{2}\left(\frac{\mathrm{d}^2 A}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}A}{\mathrm{d}r}\right) + \left(\frac{\varepsilon^2 m^2}{2r^2} + \frac{mv}{r} + \frac{\zeta}{2}\right) A = \omega A, \tag{3.42}$$

where $v = \partial_r \psi$ denotes the azimuthal velocity. There are some subtleties involved in applying WKB theory to this equation. For modes with $m > 0$, the potential diverges at the origin. This issue has long been noted in the quantum mechanics literature and can be addressed by performing a so-called Langer transform on the equation. For $m = 0$, there is no

divergence of the potential, but there is a phase shift at the origin. As pointed out by Berry and Ozorio de Almeida (1973), both cases turn out to give the same quantisation condition:

$$\frac{\sqrt{2}}{\varepsilon}\int_{r_0}^{r_1}\sqrt{\omega - V(r)}\,\mathrm{d}r = \left(n+\frac{1}{2}\right)\pi, \quad \text{with} \quad n = 0, 1, 2, \ldots, \tag{3.43}$$

where the potential is

$$V(r) = \frac{\varepsilon^2 m^2}{2r^2} + \frac{mv}{r} + \frac{\zeta}{2}. \tag{3.44}$$

If $m > 0$, the integration bounds r_0 and r_1 are the two zeros of the integrand; if $m = 0$, $r_0 = 0$ and r_1 is the one zero of the integrand. As in the case of a parallel shear flow, this expression is uniformly valid in the sense that it works for both $m \sim O(1)$ and $m \sim O(\varepsilon^{-2})$. These again correspond to anisotropic and isotropic modes, respectively, with refraction, advection, and dispersion along and across streamlines playing the same roles as before. The only difference is that the streamlines are now circular.

We consider the concrete example of an isolated Gaussian vortex on an infinite domain:

$$\psi(r) = e^{-\frac{r^2}{4}}. \tag{3.45}$$

This corresponds to an anticyclone in the centre of the domain that is surrounded by a halo of cyclonic vorticity (figure 3.6a,b). Again, the WKB calculation for $\varepsilon = \frac{1}{4}$ yields eigenvalues that agree extremely well with the exact eigenvalues (figure 3.6c). The structure of the first few modes is shown in figure 3.7. For $m = 0$, the modes are concentrated in the anticyclone. For $m > 0$, there is a repulsion from the very centre of the anticyclone due to the advection and dispersion terms in $V(r)$. This repulsion increases with m, but the modes remain primarily concentrated in the region of anticyclonic vorticity. The modes become more isotropic as m is increased. Note that again, a uniform forcing only projects onto the $m = 0$ mode due to the symmetry of the vorticity field. We also note that only the bound states form a discrete spectrum in an infinite domain, there will also be a continuum of free modes.

3.5.3 General case

Based on the intuition gained above, we wish to construct a uniformly valid asymptotic expansion for a general two-dimensional background flow. We again make a WKB ansatz that leads to (3.49). This equation can be solved by the method of characteristics and recovers Kunze's ray tracing. The quantisation condition for the general case is (3.53).

In analogy with the isotropic scaling, we begin by assuming a solution of the form

$$\hat{\phi}(x, y) = \exp\left[\frac{1}{\varepsilon^2}\sum_{j=0}^{\infty}\varepsilon^{2j}S_j(x, y)\right], \tag{3.46}$$

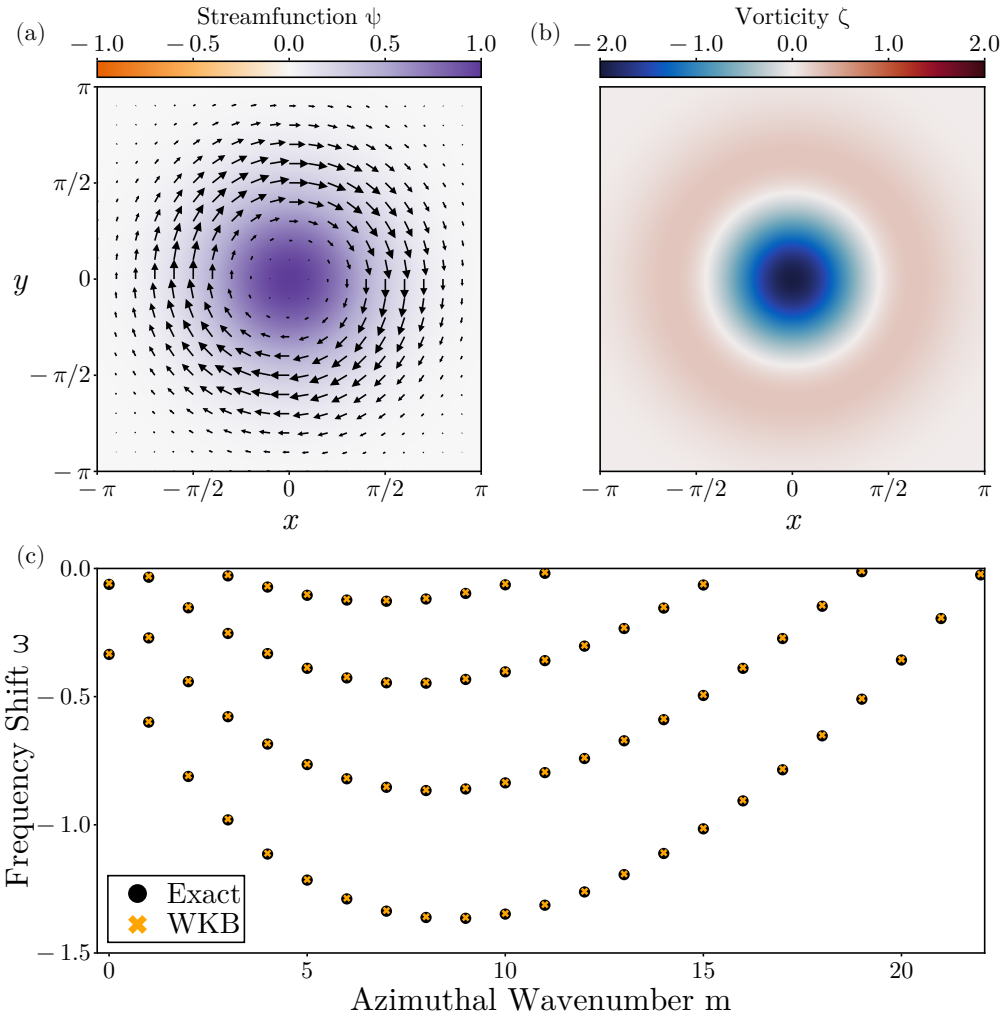

Figure 3.6: (a) Streamfunction and flow vectors for the axisymmetric flow example. (b) Vorticity field showing the anticyclonic vorticity concentrated in the centre of the domain, which is flanked by a halo of cyclonic vorticity. (c) Eigenvalues ω as a function of azimuthal wavenumber m for $\epsilon = \frac{1}{4}$. The results from WKB theory (orange crosses) are shown along with the exact eigenvalues found from numerical solutions (black circles). The WKB approximation agrees remarkably well with the numerical results.

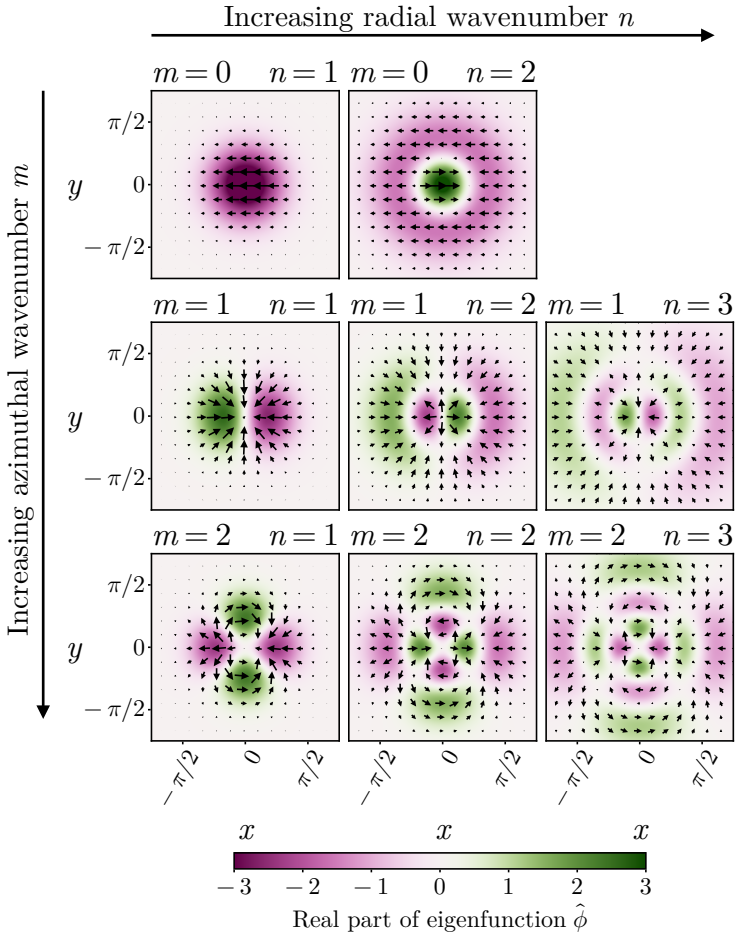

Figure 3.7: Real part of the eigenfunctions for the axisymmetric flow example with $\varepsilon = \frac{1}{4}$. The radial wavenumber *n* increases from left to right and corresponds to an increasing number of nodes in the radial direction. The azimuthal wavenumber increases from top to bottom and corresponds to an increasing number of nodes in the azimuthal direction.

again dropping the mode label. Substituting this into (3.6) yields

$$-\frac{1}{2\varepsilon^2}\left|\sum_{j=0}^{\infty}\varepsilon^{2j}\nabla S_j\right|^2-\frac{1}{2}\sum_{j=0}^{\infty}\varepsilon^{2j}\nabla^2 S_j-\frac{i}{\varepsilon^2}\sum_{j=0}^{\infty}\varepsilon^{2j}\,\mathrm{J}(\psi,S_j)+\frac{\zeta}{2}=\omega. \tag{3.47}$$

Assuming $\omega\sim O(\varepsilon^{-2})$ and collecting leading-order terms, we obtain

$$-\frac{1}{2}|\nabla S_0|^2-i\,\mathrm{J}(\psi,S_0)=\varepsilon^2\omega. \tag{3.48}$$

In the simple examples discussed above, we obtained a uniformly valid approximation by retaining the higher-order refraction term in the leading-order equation arising from an isotropic scaling. We do so again here:

$$-\frac{1}{2}|\nabla S_0|^2-i\,\mathrm{J}(\psi,S_0)+\frac{\varepsilon^2\zeta}{2}=\varepsilon^2\omega. \tag{3.49}$$

We anticipate that the order of these terms again changes for anisotropic modes. If the phase varies slowly along streamlines, the advection term is reduced by a factor $O(\varepsilon^2)$, and cross-streamline dispersion, acting on spatial variations on a scale of $O(\varepsilon)$ rather than $O(\varepsilon^2)$, will attain the same order, whereas along-streamline dispersion becomes negligible. The equation (3.49) can therefore capture both isotropic and anisotropic modes.

We now introduce the wavenumber vector $\boldsymbol{k}$ by writing $\varepsilon^{-2}\partial S_0/\partial\boldsymbol{x}=i\boldsymbol{k}$. The equation (3.49) can be solved using the method of characteristics:

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau}=\varepsilon^2\boldsymbol{k}+\boldsymbol{u},\qquad\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}\tau}=-\frac{\partial}{\partial\boldsymbol{x}}\left(\boldsymbol{u}\cdot\boldsymbol{k}+\frac{\zeta}{2}\right),\qquad\omega=\frac{\varepsilon^2|\boldsymbol{k}|^2}{2}+\boldsymbol{u}\cdot\boldsymbol{k}+\frac{\zeta}{2}. \tag{3.50}$$

These are the non-dimensionalised ray tracing equations of Kunze (1985). We further elaborate on this connection between YBJ and Kunze's ray tracing below.

Numerical solutions for the dipole flow show that the majority of a uniform forcing projects onto anisotropic modes that show little structure along streamlines and vary more rapidly across streamlines (figure 3.8). With $\varepsilon=\frac{1}{4}$ there is also some projection onto modes that show more characteristics of isotropic phase variations. The variations are more rapid, as emerges from the isotropic scaling discussed above.

Finally, we show how approximations to the eigenvalues can be obtained in the weak-dispersion limit when the flow problem is not separable, as it was in the cases of a parallel shear flow or axisymmetric flow. To this end, we utilise results from the quantum mechanics literature. Recall that the YBJ equation is equivalent to the Schrödinger equation, with the YBJ operator

$$H=-\frac{\varepsilon^2}{2}\nabla^2-i\boldsymbol{u}\cdot\nabla+\frac{\zeta}{2} \tag{3.51}$$

playing the role of the Hamiltonian. The weak-dispersion limit corresponds to the classical limit of the equivalent quantum system, and the ray tracing equations are the analogue

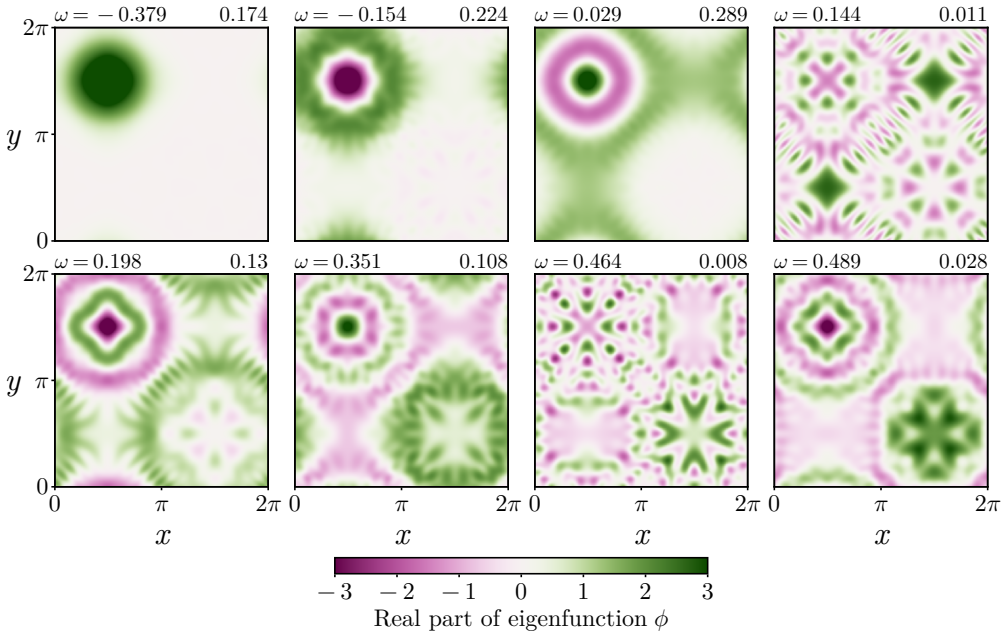

Figure 3.8: Real part of the eigenfunctions of the dipole flow with $\epsilon = \frac{1}{4}$. Together, these eight eigenfunctions represent over 97% of the energy excited by a uniform impulsive forcing. They are the eight modes with the strongest projection and are then ordered by eigenvalue ω. The eigenvalues are shown in the top left corner, and the projections of a uniform forcing onto the eigenfunction (energy fraction) are shown in the top right corner.

of the classical Hamiltonian dynamics. The classical Hamiltonian is obtained from H by making the substitution $\nabla \mapsto i\boldsymbol{k}$, yielding the dispersion relation in (3.50). The Hamiltonian dynamics are then

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\tau} = \frac{\partial\omega}{\partial\boldsymbol{k}} \quad \text{and} \quad \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}\tau} = -\frac{\partial\omega}{\partial\boldsymbol{x}}, \tag{3.52}$$

the ray tracing equations stated in (3.50). The connection with the Schrödinger equation is most easily seen in the Hamilton–Jacobi description of classical mechanics (e.g., Sakurai and Napolitano, 2020; Bühler, 2006).

The quantisation conditions derived above for separable problems, from which we obtained good approximations of the frequency shifts ω, can be generalised to some extent to non-separable problems like the dipole flow (figure 4.11). This semi-classical analysis of a quantum system was developed by Einstein (1917), Brillouin (1926), and Keller (1958), extending the Bohr–Sommerfeld quantum theory. The resulting approach is referred to as the EBK method (see also Keller, 1985; Berry and Mount, 1972; Percival, 1977). The starting point is that the rays (classical trajectories in the quantum problem), being constrained by the invariant ω (energy in the quantum problem), trace out invariant tori in the phase space spanned by $\boldsymbol{x}$ and $\boldsymbol{k}$. A ray starting on such a torus will remain on it forever. The quantisation condition selects invariant tori that correspond to allowed bound states by insisting that phase increments along closed loops on the invariant torus integrate to multiples of 2π. Recalling

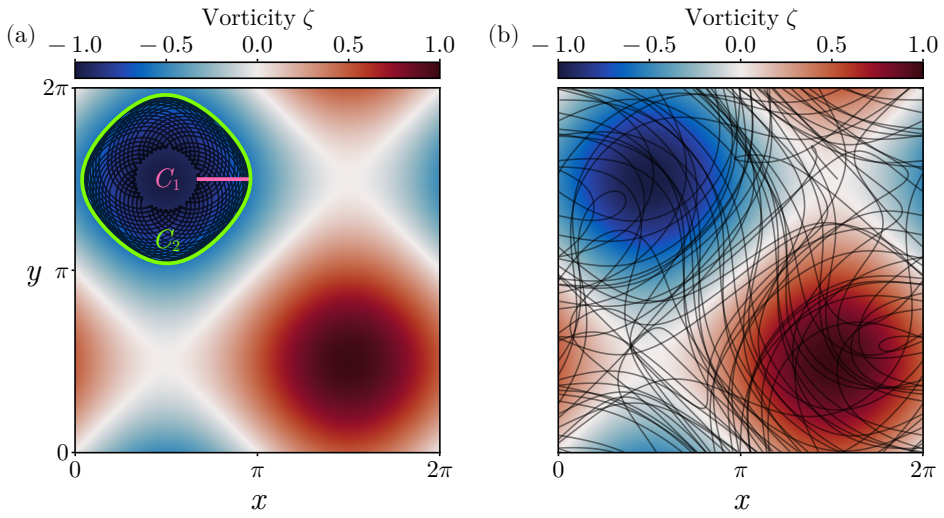

Figure 3.9: (a) Example of a trajectory tracing out an invariant torus for the dipole case. This torus corresponds to $n = 2$, $m = 0$ for $\varepsilon = \frac{1}{4}$. The background colours show the vorticity field. The black line shows a finite-time trajectory on the torus. The green and magenta lines represent a choice for the two invariant curves on the torus. They are independent because no continuous deformation of one can transform it into the other. (b) Different initial conditions result in different trajectories. This example is not bound to an invariant torus but is instead an example of a chaotic trajectory.

that $\varepsilon^{-2} S_0 = i\boldsymbol{k}$, so $\boldsymbol{k}$ is the spatial gradient of the phase, and $\boldsymbol{k} \cdot \mathrm{d}\boldsymbol{x}$ is a phase increment, the quantisation conditions read

$$\oint_{C_1} \boldsymbol{k} \cdot \mathrm{d}\boldsymbol{x} = 2\pi \left(n + \frac{1}{2}\right), \qquad \oint_{C_2} \boldsymbol{k} \cdot \mathrm{d}\boldsymbol{x} = 2\pi m, \tag{3.53}$$

where n and m are integers. The contours C_1 and C_2 are topologically independent closed curves on the invariant torus (figure 3.9a). In our example, the curve C_1 passes through the hole of the phase space torus, whereas the curve C_2 goes around the hole. The two curves are independent in the sense that neither one can be continuously deformed into the other. There is a half-integer phase shift in the quantisation condition arising from the integral along the curve C_1 because this curve passes through two caustics, the generalisation of a turning point, where additional phase shifts are incurred (Brillouin, 1926; Keller, 1958; Maslov, 1972). The curve C_2 encounters no caustics. The integer wavenumbers n and m correspond to the cross- and along-streamline variations, respectively. These EBK quantisation conditions are entirely analogous to the WKB quantisation conditions derived above for the separable parallel shear flow and axisymmetric flow.

We apply the EBK quantisation to the dipole flow with $\varepsilon = \frac{1}{4}$. Our procedure closely follows Percival and Pomphrey (1976): we find the invariant tori satisfying the quantisation condition by writing the Hamiltonian equations in action–angle variables and employing Newton's method. See Appendix 3.13 for details. All eigenvalues calculated by this EBK method show excellent agreement with the numerical values (figure 3.10).

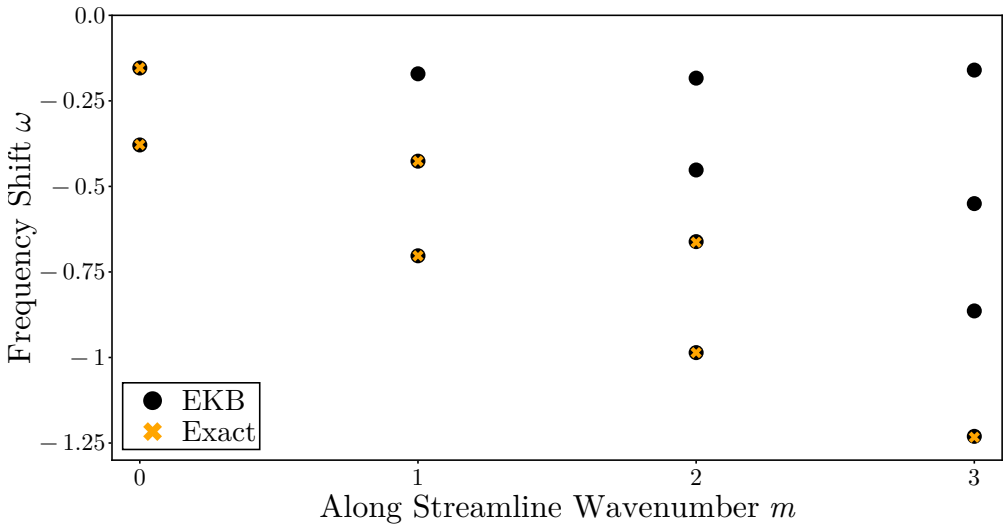

Figure 3.10: Numerical eigenvalues (black circles) and EBK eigenvalues (orange crosses) calculated for the dipole flow with $\epsilon = \frac{1}{4}$. EBK calculations are only shown for the sufficiently confined modes where the invariant tori are easy to calculate. The EBK values agree with the numerical values to $O(10^{-3})$.

As foretold by Einstein (1917), not all modes are accessible by the EBK approach. If the system is non-integrable, trajectories in phase space can become chaotic instead of tracing out an invariant torus (figure 3.9b). States corresponding to such chaotic trajectories are not amenable to the EBK method. This "quantum chaos" has received much attention in the physics literature and has connections to random matrix theory (e.g., Gutzwiller, 1992; Stone, 2005). Methods exist to estimate eigenvalues as well as their statistics (e.g., Edelman and Rao, 2005; Edelman and Sutton, 2007). We do not pursue these issues any further here, in part because a uniform forcing projects most strongly onto the regular modes accessible with the EBK method (figure 3.8).

3.6 Relation to the ray tracing equations

The previous section made clear that the ray tracing equations of Kunze (1985) are closely related to the YBJ dynamics. In the same way that Hamiltonian dynamics emerge in the classical limit of the Schrödinger equation, the ray tracing equations emerge in the weak-dispersion limit of the YBJ equation. YBJ criticised Kunze's assumption that the waves have a smaller spatial scale than the background flow, insisting that atmospheric forcing produces near-inertial waves at larger—not smaller—scales than mesoscale eddies, calling into question Kunze's ray-theoretical description in general. The analysis above clarifies that the spatial scale of the forcing is not what determines the applicability of WKB . Instead, the scale on which dynamical modes vary determines whether WKB analysis can be applied, and this spatial scale is set by how strongly dispersive the waves are. An initially uniform wave field can be thought of as consisting of a superposition of several modes, all varying on a small scale but combining into a uniform field. The distinct frequencies ω of these modes

make them dephase over time, and the superposition develops the small scales of the modes.

Our analysis also provides some additional insight into the evolution of weakly dispersive NIWs. The isotropic and anisotropic scalings show that refraction is not always of leading-order importance. The refraction term is significant only for the anisotropic modes. For isotropic modes, the refraction term is asymptotically weak and the dispersion relation is dominated by advection and dispersion. A large-scale forcing, however, will project primarily onto the anisotropic modes, as can be seen in the specific solutions for the dipole case (figure 3.8). More generally, the large values of the along-streamline wavenumber m in the isotropic case produce rapid variations that lead to strong cancellations when calculating the projection of a uniform forcing onto these modes. As such, only a weak projection can remain.

To help interpret observations from the NISKINe study, Thomas et al. (2020) performed a simplified ray tracing calculation, which predicted a rapid strain-driven growth in the wavenumber that stood in stark contrast to the data. In this region of the North Atlantic, the waves are weakly dispersive (Thomas et al., 2024a), so one may worry that this result contradicts our conclusion that ray tracing can be deployed gainfully in the weak-dispersion regime. Thomas et al. (2020) approximated the full wavevector evolution by assuming a uniform and time-independent vorticity gradient, as well as a strain field with strain rate α and its principal axis aligned with the vorticity gradient. In that setup, the wavenumber component $k_\perp$ that is aligned with the vorticity gradient, i.e., perpendicular to vorticity contours, evolves according to

$$\frac{\mathrm{d}k_\perp}{\mathrm{d}\tau} = -\frac{|\nabla\zeta|}{2} + \alpha k_\perp, \quad \text{so} \quad k_\perp = -\frac{|\nabla\zeta|}{2\alpha}\left(e^{\alpha\tau} - 1\right) \tag{3.54}$$

if $k_\perp = 0$ at time $\tau = 0$, approximating large-scale wind forcing. The exponential growth predicted by this equation does not match the data. Our analysis suggests, however, that a large-scale forcing primarily excites modes whose phase is aligned with streamlines. In this configuration, the strain is ineffective, and the initial wavenumber evolution is dominated by refraction:

$$\frac{\mathrm{d}k_\perp}{\mathrm{d}\tau} = -\frac{|\nabla\zeta|}{2}, \quad \text{so} \quad k_\perp = -\frac{|\nabla\zeta|}{2}\tau. \tag{3.55}$$

This recovers the Asselin et al. (2020) solution that Thomas et al. (2020) showed roughly matches the data. Our analysis therefore suggests that it was not ray tracing *per se* that caused the mismatch with the data but the assumptions that went into the simplified solution.[1]

Kunze (1985) considered three-dimensional ray tracing, which allows for both baroclinicity in the mean flow and a vertical wavenumber for the NIWs that corresponds to propagation

[1] It should be noted that a pure strain field does not produce a compact operator H, so the machinery based on a discrete set of eigenmodes does not apply. Instead, one should view the pure strain field as the local behaviour of some more complicated background flow that is described by a compact operator H. Because ray tracing is inherently local, the behaviour detailed above would still apply.

of the waves in the vertical. In this paper, we have restricted our attention to a barotropic mean flow and considered the propagation of a single baroclinic mode, such that the problem reduces to two-dimensional ray tracing. Exploring how the three-dimensional ray tracing is related to the full YBJ equation that also allows for baroclinicity in the background flow is left to future work.

3.7 Near-inertial wind work

One may speculate that the frequency shifts in the weak-dispersion limit could impact the energy input into NIWs by the winds. To study this, we need to consider a forced version of the YBJ equation. So far, we have focused on the problem with a horizontally uniform initial condition. This was to represent the NIW field excited by the passage of a large-scale atmospheric storm, and we studied the evolution of this NIW field in the absence of any further forcing. Real NIWs, in contrast, are continually forced by the winds, which we now represent by including a horizontally uniform forcing term in the modal YBJ equation. If we include sources of NIW energy, then we must also include the sinks, such that the wave energy can equilibrate. In the real ocean, NIW energy is primarily dissipated through mixing. Mechanisms of NIW dissipation are complicated and depend, among other factors, on the local stratification (e.g., Kunze et al., 1995; Qu et al., 2021) and the mesoscale eddy field (Sanford et al., 2021; Essink et al., 2022; Thomas et al., 2024b). For simplicity, we model these processes as a linear drag. In mixed-layer models, a linear drag is often used to model NIW propagation out of the mixed-layer (see e.g., Pollard and Millard Jr, 1970), but in our model vertical propagation is already accounted for by the decomposition into baroclinic modes. The linear drag in our model therefore represents the irreversible sink of NIW kinetic energy due to mixing. With these alterations, the modal YBJ equation reads

$$\mathrm{d}a_t = \left(-i\omega a_t - r a_t + F_t e^{ift}\right)\mathrm{d}t, \tag{3.56}$$

where a_t denotes the modal amplitude at time t, r is the linear drag coefficient, and F_t the wind forcing projected onto the mode under consideration. We suppress the mode index $\boldsymbol{\mu}$ for now but keep in mind that this equation must be solved for each mode. Note also that we have re-dimensionalised the equation here. The factor of e^{ift} back-rotates the forcing to match the back-rotated description of the NIW evolution by the YBJ equation. To proceed, we describe the wind by an Ornstein–Uhlenbeck process which satisfies

$$\mathrm{d}F_t = -c\,F_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \tag{3.57}$$

where c^{-1} is the decorrelation timescale of the wind forcing, σ is the amplitude of the stochastic excitation and W_t is a Wiener process. The power spectrum of the process F_t is

$$S(\omega) = \frac{2}{\pi}\frac{c}{c^2+\omega^2}. \tag{3.58}$$

For $\omega \gg c$ the power falls off with frequency as ω^{-2}, i.e., the spectrum is red. We find that this is a good model of the power spectrum of the wind stress from reanalysis, especially over the ocean (see Appendix 3.14 for more details).

We consider the system spun up from $t = -\infty$, such that it has statistically equilibrated for all t. This results in the formal solution for the forcing

$$F_t = \sigma \int_{-\infty}^{t} e^{-c(t-t')} \, \mathrm{d}W_{t'}, \tag{3.59}$$

and the formal solution for the mode amplitude a is given by

$$a_t = e^{-(i\omega + r)t} \int_{-\infty}^{t} e^{[i(f+\omega)+r]t'} F_{t'} \, \mathrm{d}t'. \tag{3.60}$$

The NIW kinetic-energy equation can be obtained in the usual way by multiplying (3.56) with a_t^* and adding the complex conjugate. This is allowed because it is the integral of a Wiener process that appears in (3.56), and not the Wiener process itself. The wind work Γ_t arises as

$$\Gamma_t = \frac{1}{2} \left(a_t^* e^{ift} F_t + \text{c.c.} \right). \tag{3.61}$$

We are interested in the average of Γ_t over an ensemble of many realisations of the wind-forcing. Let $\langle \cdot \rangle$ denote such the ensemble average. Hence, the ensemble average wind work is

$$\langle \Gamma_t \rangle = \frac{1}{2} \left(e^{[i(f+\omega)-r]t} \int_{-\infty}^{t} e^{[-i(f+\omega)+r]t'} \langle F_{t'}^* F_t \rangle \, \mathrm{d}t' + \text{c.c.} \right). \tag{3.62}$$

The covariance function of the Ornstein–Uhlenbeck process F_t is

$$\langle F_{t'}^* F_t \rangle = \frac{\sigma^2}{2c} e^{-c|t-t'|}, \tag{3.63}$$

so the ensemble average of Γ_t reduces to

$$\langle \Gamma_t \rangle = \frac{\sigma^2}{2c} \frac{c + r}{(c+r)^2 + (f+\omega)^2}. \tag{3.64}$$

As expected, given the initialisation at $t = -\infty$, the power input is independent of time t. This equilibrated wind work is balanced by the linear drag, such that the ensemble averaged kinetic energy in a given mode is finite. From this expression, we can furthermore see that $\langle \Gamma_t \rangle$ is smaller for $\omega > 0$ than for $\omega < 0$. This is because the wind forcing has more power at low frequencies.

We now define Q as the ratio between the equilibrium wind work in the presence of a mesoscale eddy field to the equivalent wind work in the absence of mesoscale eddies. Without mesoscale eddies, $\psi = 0$ and there are no frequency shifts, so $\omega = 0$ for the uniform mode excited by the wind. The wind work is simply

$$\langle \Gamma_t \rangle = \frac{\sigma^2}{2c} \frac{c + r}{(c+r)^2 + f^2}. \tag{3.65}$$

We calculate Q as a weighted sum of the ratio over individual modes, where the weighting is given by the projection F_{μ} of the forcing onto a given mode $\boldsymbol{\mu}$:

$$Q = \sum_{\mu} |F_{\mu}|^2 \frac{(c+r)^2 + f^2}{(c+r)^2 + (f+\omega_{\mu})^2}, \tag{3.66}$$

where we restored the subscripts for the modes. This expression depends on the wave dispersiveness ε^2 through F_{μ} and ω_{μ}. If we make the assumption that the $r \ll c$, meaning that the timescale of NIW dissipation is much longer than the memory of the winds, this expression reduces to

$$Q = \sum_{\mu} |F_{\mu}|^2 \frac{c^2 + f^2}{c^2 + (f+\omega_{\mu})^2}. \tag{3.67}$$

We use this reduced expression in the following analysis because $r \ll c$ appears reasonable and because r would be difficult to estimate.

Modulation of the NIW wind work by mesoscale eddies occurs only for $\varepsilon \lesssim 1$. Using the dipole flow as an example, we calculate Q from (3.67) as a function of c and ε (figure 3.11). For large ε, Q quickly approaches unity, regardless of the value of c. For small ε, the contours of Q become horizontal and there is little dependence of Q on ε. The dependence is primarily on c with a lower value of c resulting in a higher value of Q, i.e., a more substantial enhancement of the wind work.

Our framework provides physical motivation for why mesoscale eddies can modulate the wind work in the weak-dispersion case. Assuming $c \ll f$, which is generally the case for the wind stress over the ocean, we see that the inertial frequency f is in the ω^{-2} part of the wind power spectrum. Any process that shifts the frequency of NIWs will modulate the wind power felt by the waves. Because the wind power spectrum falls off like ω^{-2}, a shift to lower frequencies will raise the wind power felt by the waves, and a shift to higher frequencies will lower it. This is the essence of (3.67). As we have shown above, frequency shifts are small in the strong-dispersion limit, and so the waves should feel similar wind power regardless of the presence of mesoscale eddies. As such, Q is close to unity in the strong-dispersion limit. In the weak-dispersion limit, in contrast, there can be significant frequency shifts. A uniform forcing will project onto many modes with a range of frequency shifts. Due to the curvature of the wind power spectrum, going like ω^{-2}, the fractional increase in power for negative frequency shifts will be greater than the fractional decrease in power for positive frequency shifts of the same magnitude. As a result, there will be a net increase in NIW wind work when summing over all modes (see figure 3.11b for a schematic). The question remains whether this will be an appreciable effect in the ocean.

We estimate Q from observations. For each location in the ocean, we estimate ε from the deformation radius and satellite altimetry observations of the eddy field (see Appendix 3.10), and we estimate c from atmospheric reanalysis (see Appendix 3.14). We calculate the modes of the dipole flow for a range of ε, which gives us ω_{μ} and $|F_{\mu}|^2$, and we re-dimensionalise

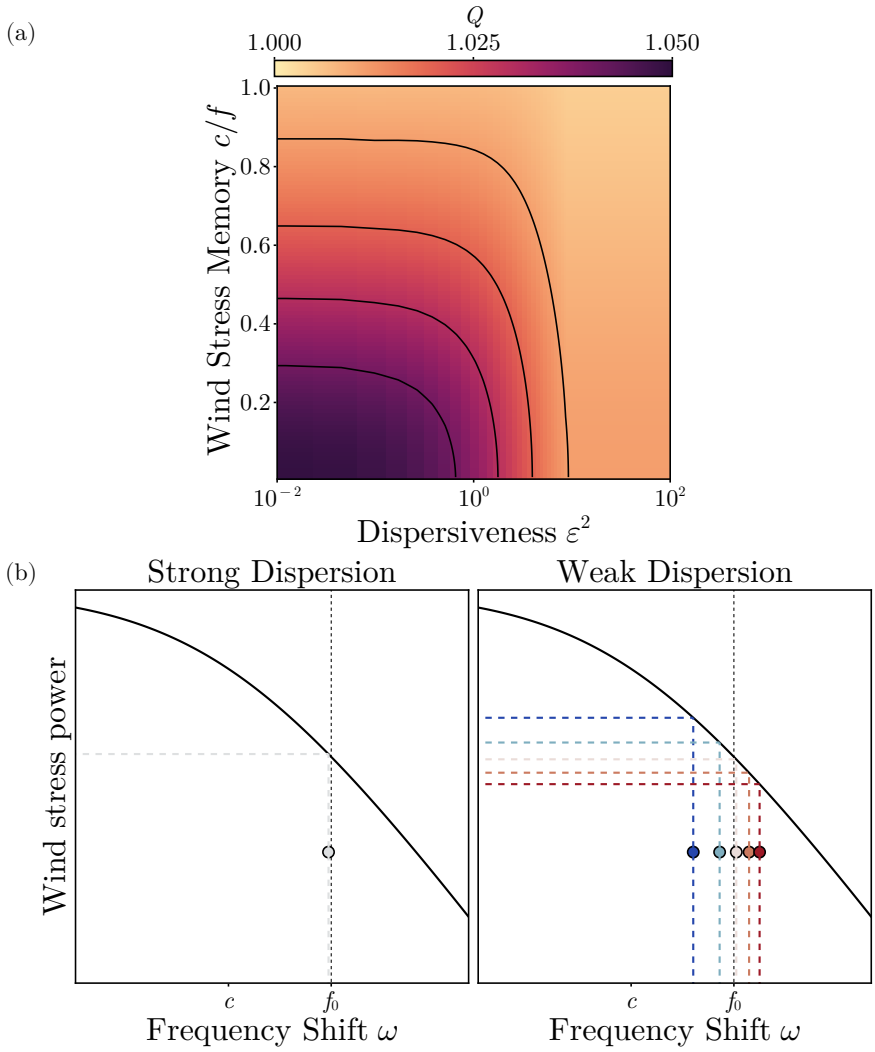

Figure 3.11: (a) Ratio Q of NIW wind work in the presence of mesoscale eddies to that without as a function of the wave dispersiveness ϵ^2 and the wind stress memory parameter c/f. Modulation of the NIW wind work by the mesoscale eddy field appears only for low ϵ and c/f. These values correspond to a re-dimensionalisation of ω by Ro $= 0.5$. (b) Schematic illustrating the enhancement of NIW wind work in the weak-dispersion regime. The solid black curve illustrates the wind stress power as a function of frequency on a log-log plot. At the inertial frequency, the power is rapidly falling off. The circles indicate the modes that a uniform initial condition projects onto. In the strong-dispersion case, the forcing projects onto a single mode with a small frequency shift. In the weak-dispersion case, the forcing projects onto a wide variety of modes with large frequency shifts.

ω_μ using the Rossby number $\mathrm{Ro} = \zeta/f$ calculated from satellite altimetry. We use the spatial structure of the vortex dipole as a stand-in for the real eddy structure. This provides an (admittedly crude) estimate of the combined effect of an anti-cyclone and a cyclone. We calculate Q by using (3.67) and then interpolating onto the correct ε.

Our estimate reveals that deviations of Q from unity are weak, at most 5%. This effect is entirely concentrated in the western boundary current regions. This is because the dimensional frequency shift scales with Ro. Over most of the ocean Ro is far too weak to produce any modulation of the NIW wind work. While this mechanism may be important for individual NIW events (Conn et al., 2024), it is clear that on average there is not a significant modulation of the NIW wind work by mesoscale eddies. The maximum modulation of 5% is significantly smaller than current uncertainties in the NIW wind work (Alford, 2020). That being said, our approximation of the wind stress as an Ornstein–Uhlenbeck process is highly simplified. The real forcing is dominated by intermittent atmospheric cyclones. The linear drag is also an extremely crude representation of NIW dissipation and should be thought of as nothing more than a stand-in for a more realistic representation.

3.8 Limitations of the model

All the limitations of the YBJ model are inherent in our analysis above. Specifically, there is no feedback of the waves onto the background flow. Xie and Vanneste (2015) extended the YBJ model by coupling it to a quasi-geostrophic model for the background flow that included wave feedbacks. These wave feedbacks can significantly alter the characteristics of the background flow (Xie and Vanneste, 2015; Wagner and Young, 2016; Rocha et al., 2018). Kafiabad et al. (2021) showed that this wave feedback can cause frequency shifts in the near-inertial waves. Furthermore, Kafiabad et al. (2021) also noticed that strong wave feedbacks can generate instabilities which cause small-scale structure in the vorticity field.

The scaling assumptions of the original YBJ equation should be kept in mind in the context of the asymptotic expansions performed above. The Rossby number is $\mathrm{Ro} = \Psi/fL^2$, and the Burger number is $\mathrm{Bu} = \lambda^2/L^2$, such that $\varepsilon^2 = \mathrm{Bu}/\mathrm{Ro}$. The YBJ equations arises asymptotically in the limit where $\mathrm{Bu} \to 0$ while Bu/Ro is kept fixed (see Asselin and Young, 2019), so $\varepsilon \ll 1$ and $\varepsilon \gg 1$ do not violate the YBJ scaling.

Thomas et al. (2017) conducted a detailed study of the evolution of NIWs in different scaling regimes. They considered a "very weak–dispersion regime" where $\mathrm{Bu} \sim \mathrm{Ro}^2$ which is equivalent to $\varepsilon^2 \sim \mathrm{Ro}$. An additional term arises compared to the YBJ equation, but they found the YBJ equation to still work well in simulations. They also considered a "strong-dispersion regime" where $\mathrm{Bu} \sim 1$. In this regime they found a leading-order uniform NIW solution, but also with the excitation of super-inertial frequencies that are not captured by YBJ. The frequency shift of the uniform mode is as predicted by YBJ. Another way to improve on the YBJ model is the YBJ$^+$ scheme of Asselin and Young (2019). It has a dispersion relation

that remains accurate over a wider range of Bu, and it has desirable numerical properties. The YBJ$^+$ equation is not a Schrödinger equation any more, however, and we do not pursue its analysis here.

Throughout this paper, we have dealt with the case in which the background flow does not evolve. In the ray tracing framework, the background flow could be allowed to evolve. The Hamiltonian operator would be time-dependent, but the equations can still be integrated along rays. For our analysis of eigenmodes to be applicable to the time-dependent case, the evolution of the background flow should be adiabatic, i.e., it should be slow compared to the wave evolution. The time for eigenmodes to dephase depends on the difference between their frequencies. In the strong-dispersion case, the frequency difference between the leading-order eigenmode and the higher eigenmodes is $O(\varepsilon^2)$, meaning that the time to dephase should be small relative to the timescale for evolution of the background flow. In the weak-dispersion limit, the eigenvalues become ever-closely packed, meaning the timescale for dephasing can become large. For the adiabatic assumption to hold, an invariant torus should deform much more slowly than the time it takes a particle to traverse the torus. If the time taken to traverse the torus is given by the advective timescale, then these two timescales are formally the same order. The adiabatic assumption will only hold if there is a symmetry which causes the torus to persist for a longer timescale. The dipole vortex is an extreme example of this where the tori never deform, yet the advective timescale is finite. In the ocean, eddies often persist as coherent features for times much longer than the advective timescale. As such, we expect that the weak-dispersion results to continue to provide insight even in the time-dependent case.

We have also assumed that the background flow is barotropic. This allows the YBJ equation to be expanded into the baroclinic normal modes. If the background flow is baroclinic, such a decomposition is not possible and the modes become coupled. This coupling of the modes means that the YBJ equation no longer reduces to the Schrödinger equation. This does not necessarily destroy the quantum analogy, as the techniques employed here may still be applicable with some modifications. We leave an exploration of these issues to future work.

Finally, we note that in the real ocean, vorticity variance increases at smaller scales. One may worry that the frequency shifts would diverge with increasing resolution. While a detailed discussion of this issue is beyond the scope of this paper, dispersion should have a regularising effect on small-scale vorticity, which leaves the problem well-posed.

3.9 Conclusions

In the YBJ framework, the evolution of NIWs in the presence of a mesoscale eddy field is governed by the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$. The limit of $\varepsilon \gg 1$ corresponds to the strong-dispersion limit and $\varepsilon \ll 1$ corresponds to the weak-dispersion limit. Both of these limits are relevant for the ocean, as the wave dispersiveness decreases with vertical mode

number and the strength of mesoscale eddies.

The YBJ equation is a Schrödinger equation, with the YBJ operator playing the role of the Hamiltonian operator in quantum mechanics. As is conventional in quantum mechanics, the evolution of NIWs can be described using the eigenmodes of the YBJ operator and their eigenvalues, which determine the frequency shift away from the inertial frequency. Perturbation methods from quantum mechanics yield insight into YBJ dynamics and its relationship to the ray tracing equations of Kunze (1985).

In the strong-dispersion regime $\varepsilon \gg 1$, perturbation theory yields closed-form expressions for the NIW modes. To leading order, a spatially uniform forcing excites a spatially uniform NIW mode. This mode is modulated by an order ε^{-2} perturbation proportional to the streamfunction of the eddy field. The frequency shift is also of order ε^{-2} and proportional to the average kinetic energy of the eddies. Both of these results recover predictions from Young and Ben-Jelloul (1997) through an alternative approach. The same approach also yields expressions for the modes that are not spatially uniform to leading order. The degeneracy of these modes at leading order is lifted at higher order, and the frequency shifts and spatial structures can be determined. Wind patterns associated with sharp atmospheric fronts may excite these modes more strongly than the uniform forcing assumed throughout this work (e.g., Thomas, 2017).

In the weak-dispersion regime $\varepsilon \ll 1$, the YBJ equation is amenable to WKB analysis. In simple (separable) background flow geometries, this allows the straightforward calculation of eigenmodes and their frequency shifts, which are excellent approximations of the exact frequency shifts even for modestly small ε. More generally, the weak-dispersion limit of the YBJ equation corresponds to the classical limit of quantum mechanics. The YBJ equation reduces to the ray equations of Kunze (1985), the equivalent to the corresponding classical Hamiltonian dynamics. The semi-classical EBK analysis allows the calculation of frequency shift for non-separable background flows for the regular part of the spectrum, which again are in excellent agreement with the full shifts. The emergence of the ray equations in the classical limit furthermore suggests that they can be applied if dispersion is weak, whether or not the forcing has a large horizontal scale. The spatial-scale separation underlying the ray equations emerges because a uniform initial condition projects onto many modes, and these modes exhibit small-scale structure.

The frequency shift of NIW modes away from the inertial frequency implies that the NIW wind work can be modulated by mesoscale eddies. We quantify this using Q which measures the ratio of the NIW wind work in the presence of mesoscale eddies to that without mesoscale eddies. This modulation arises due to the curvature of the wind power spectrum, which enhances the power input into modes with a shift to lower frequencies more than it suppresses the power input into modes with a shift to higher frequencies. On average, this effect is weak in the ocean, however, with the modulations always being less than 5%.

Acknowledgements The authors thank three anonymous reviewers for their insightful comments on the manuscript, and they gratefully acknowledge support from NASA under grants 80NSSC22K1445 and 80NSSC23K0345, from NSF under grant OCE-1924354, and from the Simons Foundation Pivot Fellowship program.

Data Availability Statement Code to numerically solve the 2D eigenvalue problem is available at `https://github.com/joernc/ybjmodes`. The SSH data is available from the E. U.'s Copernicus Marine Service at `https://doi.org/10.48670/moi-00148`. The ERA5 reanalysis data is available from the Copernicus Climate Change Service (C3S) Climate Data Store at `https://doi.org/0.24381/cds.adbb2d47`. The ECCO density data is available from `https://doi.org/10.5067/ECG5D-ODE44`.

Declaration of Interests The authors report no conflict of interest.

3.10 Appendix A: Calculating the wave dispersiveness

Here, we describe the calculations used to estimate the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$ from observations. At each location, we estimate the set of deformation radii λ from hydrography and the characteristic strength of the streamfunction Ψ from altimetry.

Following Smith (2007), we calculate λ by solving the baroclinic eigenvalue equation using finite differences. We perform this calculation using the climatology from the Estimating the Circulation and Climate of the Ocean (ECCO) state estimate version 4 release 4 (Fukumori et al., 2020; Forget et al., 2015). We solve the baroclinic eigenvalue equation at each horizontal grid cell on the ECCO grid to obtain maps of the deformation radii. We display ε^2 for the lowest four baroclinic modes only, for which the numerical approximation has a minimal effect.

To calculate Ψ, we use sea surface height (SSH) observations from the Data Unification and Altimeter Combination System's (DUACS) delayed-time (DT) 2018 release Taburet et al. (2019). The SSH is provided at a (nominal) $(1/4)^\circ$ and daily resolution. We calculate a geostrophic streamfunction using $\psi = g\eta/f$, where η is the SSH and f is the (now latitude-dependent) Coriolis frequency. We take observations from 2007 to 2022 and estimate Ψ as the RMS ψ over that period. Again, we are assuming that the streamfunction is barotropic.

3.11 Appendix B: Numerical solutions to the eigenvalue problem

To numerically solve the eigenvalue problem (3.6), we discretise the Hamiltonian operator H using finite differences. We use a fourth-order central difference scheme for the Laplacian operator in the dispersion term. For the advection term, we employ the fourth-order enstrophy-conserving scheme of Arakawa (1966), which preserves the Hermitian nature of the operator and translates into energy conservation in this context. In the notation of Arakawa (1966), we employ $2J_1 - J_2$, where $J_1 = \frac{1}{3}(J^{++} + J^{+\times} + J^{\times+})$ and $J_2 = \frac{1}{3}(J^{\times\times} + J^{\times+} + J^{+\times})$. For the refraction term, we evaluate ζ analytically at each point,

although in general it could be calculated from the streamfunction using finite differences as well.

We use a spatial resolution of up to 1024×1024 points and solve for the lowest eigenvalues using Lanczos iteration. The resolution is chosen by checking the convergence of the eigenvalues. The number of eigenvalues solved for depends on the value of ε, which controls how densely packed the eigenvalues are and thus how many must be computed to find all eigenmodes that a uniform forcing projects onto substantially. We ensure a large enough number of eigenvalues are computed by summing the square of the projection coefficients.

3.12 Appendix C: Analytical solutions to shear flow WKB integrals

Here, we provide analytical solutions to the WKB problem with the sinusoidal shear flow. First we rewrite the potential as

$$V(x) = A_m \cos(x + \delta_m) + \frac{\varepsilon^2 m^2}{2}, \quad \text{where} \quad A_m = \sqrt{m^2 + \frac{1}{4}} \quad \text{and} \quad \tan \delta_m = -2m. \tag{3.68}$$

If $m > 0$ and arctan corresponds to the principal value, then it follows that $\delta_m = \pi + \arctan(-2m)$. Because the domain is periodic, we can consider any interval of length 2π. For convenience, we choose $[-\pi - \arctan(-2m), \pi - \arctan(-2m)]$. We can now make the change of variable $x' = x + \arctan(-2m)$. The transformed potential is

$$V(x') = \frac{\varepsilon^2 m^2}{2} - A_m \cos(x'). \tag{3.69}$$

With the potential in this form, the WKB integral (3.32) can be evaluated in terms of the elliptic integral of the second kind $E(\varphi|k^2)$:

$$S_0 = \pm 2\sqrt{2} i \varepsilon \sqrt{\omega - \frac{\varepsilon^2 m^2}{2} + A_m}\; E\left(\frac{x'}{2}\middle| \frac{2A_m}{\omega - \frac{\varepsilon^2 m^2}{2} + A_m}\right). \tag{3.70}$$

We can obtain an equation for the eigenvalues from (3.39). Letting x_1' denote the positive turning point given by

$$x_1' = \pi - \arccos\left(\frac{\omega - \frac{\varepsilon^2 m^2}{2}}{A_m}\right), \tag{3.71}$$

we obtain

$$E(\varphi|k^2) = \frac{\varepsilon \pi \left(n + \frac{1}{2}\right)}{4\sqrt{2\left(\omega - \frac{\varepsilon^2 m^2}{2} + A_m\right)}}, \quad \text{where} \quad \varphi = \frac{x_1'}{2} \quad \text{and} \quad k^2 = \frac{2A_m}{\omega - \varepsilon^2 m^2/2 + A_m}. \tag{3.72}$$

This is a transcendental equation that can be solved numerically for the eigenvalues ω.

The eigenvectors can be normalised by requiring

$$\int_{-\pi}^{\pi} [\hat{\phi}_n(x)]^2 dx = 2\pi. \tag{3.73}$$

Letting C be the normalisation constant, we obtain

$$\frac{8C^2}{\sqrt{\omega - \frac{\varepsilon^2 m^2}{2} + A_m}} F(\varphi|k^2) = 2\pi, \tag{3.74}$$

where $F(\varphi|k^2)$ is the elliptic integral of the first kind (see Bender and Orszag, 1999).

The projection of a uniform forcing onto a given mode with even symmetry about the bottom of the potential is

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\phi}_n(x) dx. \tag{3.75}$$

This integral can be evaluated (again see Bender and Orszag, 1999) as

$$|a_n|^2 = \frac{1}{\sqrt{A_m - \omega + \frac{\varepsilon^2 m^2}{2}}} \frac{\varepsilon}{\sqrt{2} F(\varphi|k^2)}. \tag{3.76}$$

3.13 Appendix D: Further Details about the EBK Method

We principally follow Percival and Pomphrey (1976) to calculate the invariant tori satisfying quantisation conditions and the associated EBK predictions for the eigenvalues ω. We write the angle Hamilton equations, which are partial differential equations that describe the invariant torus:

$$\nu \frac{\partial x}{\partial \theta} + \mu \frac{\partial x}{\partial \varphi} = \varepsilon^2 k + u, \tag{3.77}$$

$$\nu \frac{\partial y}{\partial \theta} + \mu \frac{\partial y}{\partial \varphi} = \varepsilon^2 l + v, \tag{3.78}$$

$$\nu \frac{\partial k}{\partial \theta} + \mu \frac{\partial k}{\partial \varphi} = -\left(k \frac{\partial u}{\partial x} + l \frac{\partial v}{\partial x} + \frac{1}{2} \frac{\partial \zeta}{\partial x} \right), \tag{3.79}$$

$$\nu \frac{\partial l}{\partial \theta} + \mu \frac{\partial l}{\partial \varphi} = -\left(k \frac{\partial u}{\partial y} + l \frac{\partial v}{\partial y} + \frac{1}{2} \frac{\partial \zeta}{\partial y} \right). \tag{3.80}$$

The quantization conditions can then be written as integrals over the angles θ and φ:

$$\int \left(k \frac{\partial x}{\partial \theta} + l \frac{\partial y}{\partial \theta} \right) \mathrm{d}\theta = 2\pi m, \qquad \int \left(k \frac{\partial x}{\partial \varphi} + l \frac{\partial y}{\partial \varphi} \right) \mathrm{d}\varphi = 2\pi(2n + 1 + m). \tag{3.81}$$

The integration along θ passes around the hole of the torus (like contour C_2 in figure 3.9). The integration along φ passes through the hole twice and also around the hole once, so we double the radial phase increment $2\pi(n + \frac{1}{2})$ and add the azimuthal phase increment $2\pi m$ in the second quantization condition. We average these numerical integrals over the respective other coordinate to increase the accuracy.

We discretise the above equations by dividing the $[0, 2\pi]$ intervals that the angles θ and φ vary over using 64 points and approximate derivatives using an eighth-order finite difference scheme. We initialise the calculation with $\theta = -\frac{1}{2}$, $\varphi = \frac{1}{10}$,

$$x = \frac{1}{2}(1+\cos\varphi)\cos\theta, \quad y = \frac{1}{2}(1+\cos\varphi)\sin\theta, \quad k = -\varepsilon^{-1}\sin\varphi\cos\theta, \quad l = -\varepsilon^{-1}\sin\varphi\sin\theta \tag{3.82}$$

for the $(n, m) = (0, 0)$ torus and apply Newton iteration to satisfy the above equations. We determine ω by applying the dispersion relation at each point of the torus and averaging over all grid points. We then change the quantum numbers to other values and start the Newton iteration from the previously found torus, using iterations at intermediate values if needed.

3.14 Appendix E: Estimating the decorrelation time of wind stress

Here, we describe the calculations used to estimate the decorrelation time c^{-1} of the wind stress. For the wind forcing, we use the European Centre for Medium-Range Weather Forecasting ERA-5 reanalysis (Hersbach et al., 2018). For the calculations below, we use data from 2015 to 2020. At each grid cell, we use the 10 m zonal (u_w) and meridional (v_w) winds with hourly resolution. Following Pollard and Millard Jr (1970) we convert this to a wind stress using a bulk aerodynamic drag formulation. The time series at each location is used to calculate a power spectrum of the wind stress. The decorrelation timescale is obtained by fitting the following model to the estimated spectrum:

$$S(\omega) = \frac{A}{1 + \left(\frac{\omega}{c}\right)^s}, \tag{3.83}$$

where A and s are additional fitted parameters that we do not use here. Over the ocean $s = 2$ is a reasonable approximation, which motivates our use of the Ornstein–Uhlenbeck process above.

Chapter 4

GLOBAL NEAR-INERTIAL WAVE SPECTRA SHAPED BY MESOSCALE EDDIES

This chapter has been submitted to the *Journal of Physical Oceanography*. As a self-contained work, some notation may differ from conventions used elsewhere in this thesis.

4.1 Abstract

Wind-forced near-inertial waves (NIWs) propagate through a sea of mesoscale eddies, which can fundamentally alter their evolution. The nature of this NIW–mesoscale interaction depends on how dispersive the waves are. For weakly dispersive waves, ray tracing suggests that the NIW frequency should be shifted by $\frac{1}{2}\zeta$, where ζ is the mesoscale vorticity, and that the waves are refracted into anticyclones. Strongly dispersive waves, in contrast, retain the large-scale structure of the wind forcing and exhibit a small negative frequency shift. Previous in situ observational studies have indeed revealed varying degrees of NIW–mesoscale interaction. Here, observations of NIWs from drifters are used to map the geography of NIW–mesoscale interactions globally, and idealized simulations and a simple model are used to identify the underlying physical processes. Almost everywhere in the ocean, with the notable exception of the Northeast Pacific, the NIW frequency is strongly modulated by the mesoscale vorticity, with the slope of the frequency shift vs. vorticity taking values of approximately 0.4. Concentration of NIW energy into anticyclones is a common feature throughout the ocean. Other aspects of the observations, however, show signatures of strongly dispersive waves: a negative frequency shift and weaker concentration into anticyclones in regions with strong eddies as well as weak modulation of the NIW frequency by mesoscale eddies in the Northeast Pacific. The signatures of both weakly and strongly dispersive NIW behavior can be rationalized by the geography of the wave dispersiveness and the fact that wind forcing excites multiple vertical modes with different wave dispersiveness.

4.2 Introduction

Near-inertial waves (NIWs) excited by atmospheric storms substantially influence upper-ocean dynamics through their role in small-scale mixing (e.g., Alford et al., 2016). They are characterized by strong vertical shear that can trigger shear instabilities, causing turbulence and mixing. This process deepens the surface mixed-layer in the aftermath of a storm, and a significant fraction of NIW energy is dissipated in the upper ocean (Alford, 2001; Plueddemann and Farrar, 2006; Alford, 2020). Ultimately, NIW-induced mixing influences the ocean's surface heat budget, impacting atmospheric circulation and precipitation patterns (Jochum et al., 2013).

The evolution of NIWs can be strongly influenced by interactions with mesoscale eddies. If the waves are weakly dispersive, they can be described by ray tracing (Kunze, 1985; Conn et al., 2025), and one expects them to be refracted toward anticyclonic vorticity. There is now substantial observational evidence for such ζ refraction and the resulting concentration of NIW energy into anticyclones (Elipot et al., 2010; Thomas et al., 2020; Yu et al., 2022b; Conn et al., 2024). This refraction also increases the speed at which NIWs propagate to depth (Lee and Niiler, 1998; Asselin and Young, 2020), which in turn impacts where in the upper ocean NIWs generate mixing (e.g., Essink et al., 2022; Alford et al., 2025).

The recent evidence of ζ refraction from field campaigns in the North Atlantic (Thomas et al., 2020; Yu et al., 2022a; Conn et al., 2024) stands in contrast to the pioneering Ocean Storms Experiment in the Northeast Pacific (D'Asaro et al., 1995), where the NIW–mesoscale interaction was found to be weak. Studying the scale reduction of the NIW field after the passage of a storm, D'Asaro et al. (1995) could explain the evolution as driven by the β effect, so-called β refraction (Munk and Phillips, 1968; Gill, 1984), with mesoscale eddies playing no discernible role (D'Asaro, 1995). Young and Ben-Jelloul (1997, hereafter YBJ) argued that the lack of NIW–mesoscale interaction during Ocean Storms was due to a breakdown of the scale separation assumption of ray tracing, showing that strongly dispersive waves remain largely uniform in the presence of eddies. Indeed, Thomas et al. (2024a) showed that the differences in NIW behavior between Ocean Storms and the NISKINe experiment in the North Atlantic could be attributed to differences in how dispersive the waves were in these two regions.

Given these drastic differences in NIW behavior, the goal of this paper is to characterize the geography of NIW–mesoscale interaction globally. We use YBJ's description of the NIW evolution in the presence of a mesoscale eddy field both as a guide in the analysis and as our main interpretive framework. Throughout the paper, we consider the YBJ equation for a single vertical mode, which requires assuming a barotropic eddy field. For a vertical mode with structure $g(z)$, the NIW velocity can be expressed as $(u_w(x,y,t), v_w(x,y,t))g(z)$. The YBJ equation is then formulated as an evolution equation for $\phi = (u_w + iv_w)e^{ift}$, where f is the inertial frequency and here assumed constant (i.e., we neglect the β effect). Under these assumptions, the YBJ equation can be written as

$$\frac{\partial \phi}{\partial t} + \mathrm{J}(\psi, \phi) + \frac{i\zeta}{2}\phi - \frac{if\lambda^2}{2}\nabla^2\phi = 0, \tag{4.1}$$

where ψ is the prescribed streamfunction of mesoscale eddies, $\zeta = \nabla^2\psi$ is the associated vorticity, λ is the deformation radius of the vertical mode under consideration, and $\mathrm{J}(a,b) = \partial_x a\,\partial_y b - \partial_y a\,\partial_x b$ is the Jacobian operator. The second term in (4.1) represents the advection of NIWs by mesoscale eddies, the third term represents refraction by mesoscale vorticity, and the fourth term represents dispersion. The relative importance of dispersion versus refraction is captured by the wave dispersiveness $\varepsilon^2 = f\lambda^2/\Psi$, where Ψ is an appropriate scale for the mesoscale streamfunction.

The wave dispersiveness ε^2 governs the extent to which NIW–mesoscale interactions influence the evolution of the waves (Fig. 4.1). When $\varepsilon \gg 1$, dispersion dominates and mesoscale effects are limited. When $\varepsilon \ll 1$, dispersion is weak and the mesoscale strongly modulates NIW behavior. Kunze (1985) used ray-tracing to describe how NIWs interact with mesoscale eddies, proposing that the eddies act to shift the local inertial frequency to an effective value $f + \zeta/2$. This ray-tracing framework applies in the weak-dispersion limit, $\varepsilon \ll 1$, when the scale separation assumption is appropriate, even when the atmospheric forcing is large-scale (Conn et al., 2025), providing a useful prediction for how mesoscale eddies should modulate the NIW frequency. This frequency shift also sets up phase gradients in the NIWs. Dispersion, which is weak but non-zero, acts to flux energy down these phase gradients, resulting in the concentration of NIW energy into anticyclones, although advective straining can complicate this process (Rocha et al., 2018; Conn et al., 2025). If dispersion is strong, in contrast, the waves only take on weak horizontal structure that is proportional to the streamfunction ψ and suppressed by a factor of order ε^{-2} compared to the leading-order uniform structure (Young and Ben-Jelloul, 1997; Conn et al., 2025). Similarly, the NIW frequency in this regime is negatively shifted by a factor of order ε^{-2} compared to the scale of ζ and proportional to the area-averaged mesoscale kinetic energy rather than ζ (Young and Ben-Jelloul, 1997; Conn et al., 2025):

$$\Delta\omega = -\frac{1}{f\lambda^2}\frac{\int \frac{1}{2}|\nabla\psi|^2 \,\mathrm{d}^2x}{\int \mathrm{d}^2x}. \tag{4.2}$$

Conn et al. (2025) showed that a horizontally uniform wind forcing primarily projects onto a single mode, meaning the spectrum of the NIW should be strongly peaked in this strong-dispersion case.

The global availability of surface drifters and satellite altimetry provides an opportunity to assess the importance of NIW–mesoscale interactions across the world ocean. In situ field campaigns offer detailed spatial information over limited regions but cannot achieve global coverage. Altimetry by itself is blind to NIWs, as they produce no leading-order pressure signal, but it can be used to characterize the mesoscale eddy field. Drifters often exhibit easily discernible inertial circles (e.g., D'Asaro et al., 1995; ber), and frequency spectra universally exhibit a distinct near-inertial peak (Yu et al., 2019). The drifter observations' Lagrangian nature means that they capture the waves' intrinsic frequencies and are not affected by Doppler shifts.

Elipot et al. (2010) previously characterized the global spectral properties of NIWs using drifter data. Averaged globally, they found that the frequency of the NIW peak roughly followed 0.4ζ, close to the $\zeta/2$ prediction from Kunze (1985). They identified wind forcing as the primary driver of NIW amplitude variations but also found ζ to modulate the amplitude, consistent with ray tracing and YBJ theory. Elipot et al. (2010) further linked variations in the width of the NIW peak to the Laplacian of ζ, based on theoretical arguments by Klein

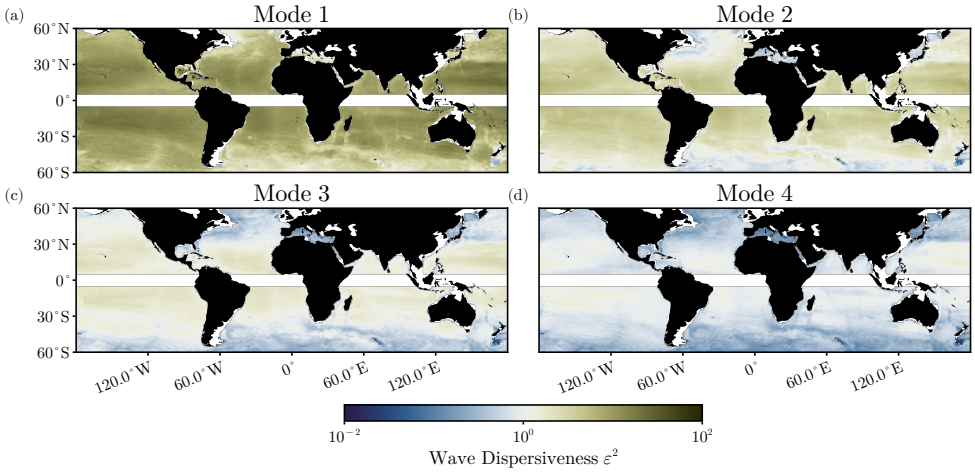

Figure 4.1: Geography of the wave dispersiveness ε^2 of the first four vertical modes, showing a transition from strongly dispersive low modes to more weakly dispersive high modes. Note the weaker dispersion in major current systems. Figure adapted from Conn et al. (2025). These maps represent annual averages, with the streamfunction magnitude Ψ coming from altimetry measurements (Taburet et al., 2019) and the deformation radius λ being calculated from the climatological density profiles from the Estimating the Circulation and Climate of the Ocean (ECCO) state estimate version 4 release 4 (Fukumori et al., 2020; Forget et al., 2015).

et al. (2004). Park et al. (2005) also studied the amplitude of NIWs globally using the drifter dataset, while Park et al. (2009) used the YBJ model in the absence of a background flow (but retaining β refraction) to investigate the decay timescales of NIWs in the drifter data.

The strength of mesoscale eddies, the inertial frequency, the deformation radius, and the projection of wind forcing onto vertical modes all vary substantially across the world ocean. Because these quantities shape the characteristics of NIW–mesoscale interactions, one should expect substantial regional differences as exemplified by the dichotomy between Ocean Storms and NISKINe (Thomas et al., 2024a). These regional differences can be obscured in a global mean, so we leverage the expanded drifter dataset available since Elipot et al. (2010) to explore this geography.

This analysis shows that mesoscale eddies modulate the NIW frequency nearly everywhere, with the notable exception of the Northeast Pacific. In highly energetic regions, we also find evidence for the excitation of strongly dispersive waves. We interpret these results emphasizing the role of the wave dispersiveness in modulating NIW–mesoscale interactions. Idealized simulations and a simple model of this interaction reproduce key spectral features observed in the drifter data. Together, these results show how NIW–mesoscale interactions have a major organizing influence on the global structure of near-inertial energy.

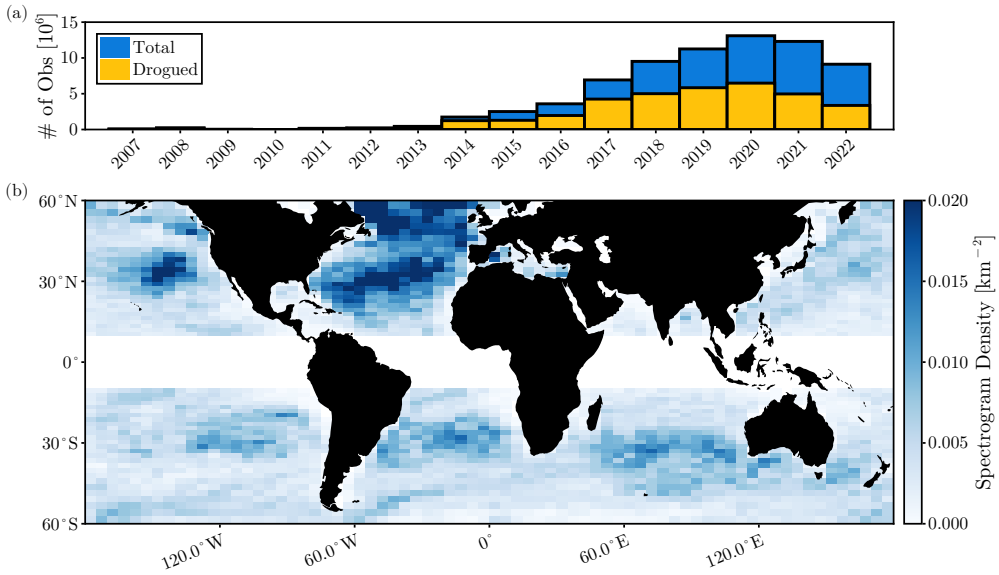

Figure 4.2: Availability of drifter data. (a) Number of hourly drifter observations available per year. Blue refers to the full set of GPS-tracked drifters, while yellow represents only those observations for which the drifter's drogue is still attached. (b) Distribution of drifter observations, shown as the number of spectrograms per unit area in a given grid cell. The pattern reflects the large-scale convergence and divergence in the ocean as well as a northern-hemisphere and Atlantic bias.

4.3 Drifter Observations of NIWs

To characterize NIWs globally, we use observations from the Global Drifter Program. This dataset provides hourly records of both position and horizontal velocity for 19 396 drifters spanning the period 2007–2023. We restrict our analysis to drifters with GPS-tracked positions. Each drifter is initially drogued, such that its velocity reflects the current at 15 m depth; however, the drogue can be lost over the course of a drifter's lifetime. Observations made while the drogue is still attached account for approximately 49% of the dataset (Fig. 4.2a). We discard all measurements following drogue loss. The distribution of drifters across the ocean is non-uniform and reflects large-scale patterns of horizontal convergence and divergence as well as preferential deployment, for example in the North Atlantic (Fig. 4.2b).

To calculate NIW spectra, we first divide the drifter trajectories into overlapping 20-day segments and compute the spectrogram of $u + iv$ for each segment. These spectrograms are calculated using a Lanczos window and subsequently binned based on the mean position of each segment, also determined using the same window. A defining feature of NIWs is their circular polarization. In the Northern Hemisphere (where $f > 0$), NIWs exhibit clockwise polarization, so we retain only the clockwise component of the power spectrum; in the Southern Hemisphere (where $f < 0$), we retain only the counterclockwise component. The 20-day segment results in a spectral resolution of 3.6×10^{-6} s^{-1}. Spectral estimates are calculated as averages over geographical bins (bins with less than 10 such 20-day segments are dismissed). The grid spacing of the binning is 5° zonally and variable in the meridional

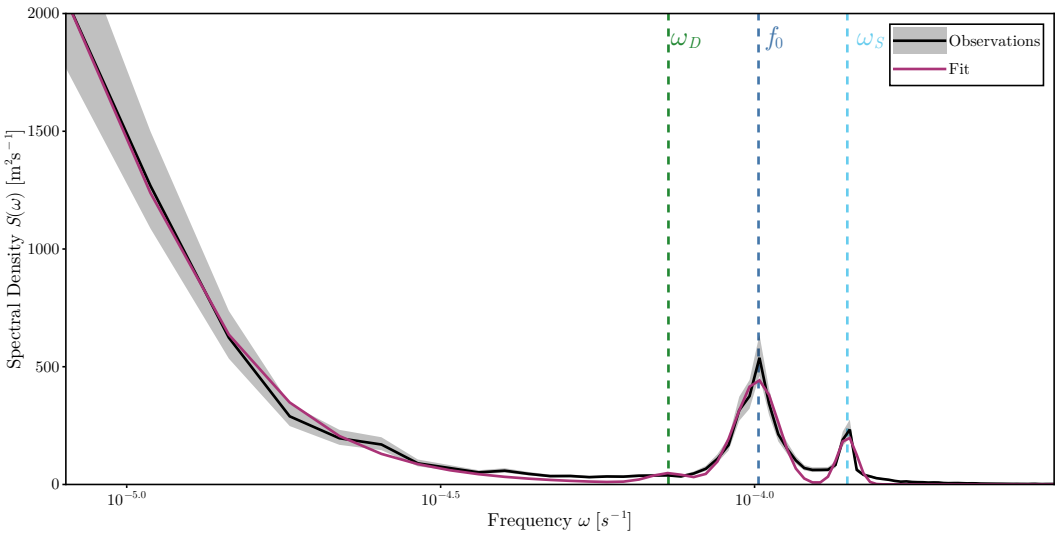

Figure 4.3: Example spectral estimate (black, with 95% confidence interval in gray) and the associated least-squares fit of the model spectrum (purple). This example shows peaks corresponding to NIWs as well as the semidiurnal tide. There is little power associated with the diurnal tide in this example. The fitted model has a low-frequency component with a slope of -3.13, a NIW peak at $1.00 f_0$ with a width of $0.08 f_0$, and a semidiurnal peak at $1/38 f_0$ with a width of $0.05 f_0$.

direction. The meridional grid spacing is capped at 5° and decreases toward the equator to ensure that the variation in f across each grid cell remains smaller than the spectral resolution, rendering the impact of f variations on the spectral estimates negligible.

The spectra show evidence of low-frequency balanced motion, NIWs, diurnal/semi-diurnal tides, and high-frequency internal waves. We perform a spectral fit to isolate the NIW signal. The model spectrum consists of a low-frequency component and Gaussian peaks for the NIWs and tides. We perform the fit in linear space, and we do not include a term for the internal wave continuum as its amplitude is orders of magnitude smaller than the other terms. The spectral model is

$$S(\omega) = \frac{A_L}{1 + (\frac{\omega}{\omega_L})^s} + A_I e^{-\frac{(\omega - f - \omega_I)^2}{2\sigma_I^2}} + A_D e^{-\frac{(\omega - \omega_D)^2}{2\sigma_D^2}} + A_S e^{-\frac{(\omega - \omega_S)^2}{2\sigma_S^2}}, \tag{4.3}$$

where A_L, A_I, A_D, A_S are the amplitudes of the low-frequency motion, NIWs, diurnal tides, and semi-diurnal tides, respectively, ω_L is the transition frequency of the low-frequency model, s is the spectral slope of the low-frequency model, ω_I is the NIW frequency shift, σ_I is the NIW spectral width, ω_D is the frequency of the diurnal tides, σ_D is the width of the peak at ω_D, ω_S is the frequency of the semi-diurnal tides, and σ_S is the width of the peak at ω_S. As we consider either the clockwise or counterclockwise components of the power spectrum alone, we take the frequencies to be positive. With this convention $\omega_I > 0$ corresponds to a shift of the NIWs to frequencies higher than f (i.e., superinertial frequencies), while $\omega_I < 0$ corresponds to a shift to frequencies lower than f (i.e., subinertial frequencies).

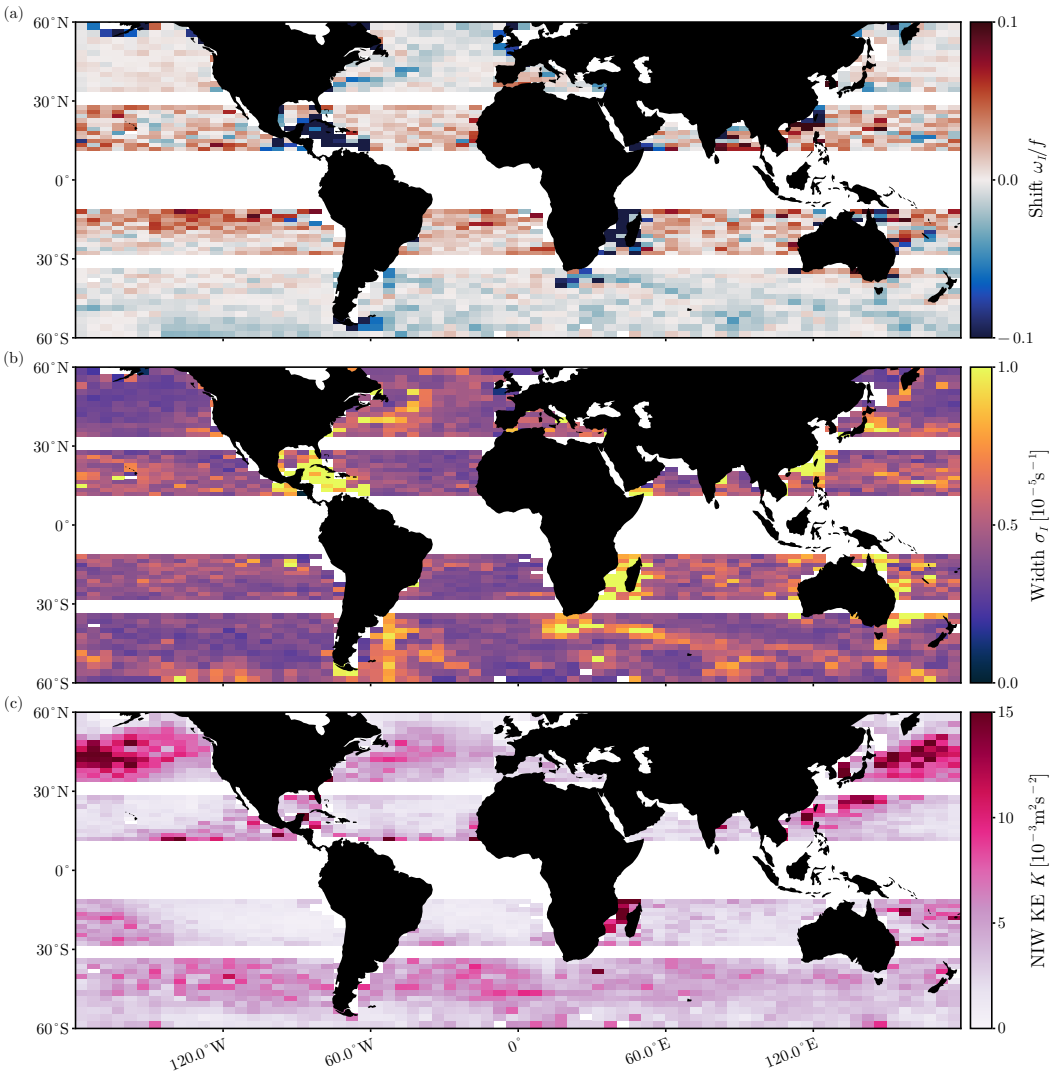

Figure 4.4: Characteristics of the NIW peak: (a) the frequency shift ω_I/f, (b) the spectral width σ_I, and (c) the NIW kinetic energy K. All estimates are calculated from averages over all available spectrograms in a given geographical bin. 95% confidence intervals are shown in Suppl. Fig. 1.

Given a spectral estimate $\hat{S}$, we fit the model spectrum to this estimate using the least squares method (Appendix A). We fix the tidal frequencies and determine all other parameters through the fit (see Fig. 4.3 for an example). The NIW kinetic energy can be obtained from the fit by integrating the NIW part of the model spectrum across all frequencies:

$$K = \sqrt{2\pi} A_I \sigma_I. \tag{4.4}$$

Each spectral estimate $\hat{S}$ is obtained by averaging over all available spectrograms in the respective bin. The spectral estimate $\hat{S}$ follows a χ^2 distribution, where the number of degrees of freedom is equal to twice the number of spectrograms that are averaged to obtain $\hat{S}$. We draw random samples from this distribution and perform the fitting procedure on each sample to obtain an empirical distribution for the fitted parameters.

First, we apply the fitting procedure to spectral estimates obtained by averaging all spectra within a given spatial bin. We exclude regions near the equator, where NIWs merge with low-frequency tropical wave modes, and around the turning latitudes of the diurnal tides, where NIWs cannot be distinguished from tidal signals. In these regions, the spectral diagnostics are dominated by changes in the relative position of the tidal frequency and f, with any mesoscale signal not being observable. Equatorward of 30°, the mean frequency shift ω_I is generally positive (Fig. 4.4a). Elipot et al. (2010) attributed this to the equatorward propagation of NIWs—a plausible explanation that we revisit in the discussion (Section 4.6).

Poleward of 30°, we observe regions with both positive and negative frequency shifts. Negative shifts occur primarily in energetic regions such as western boundary currents and the Antarctic Circumpolar Current (ACC). We hypothesize that these shifts result from the excitation of strongly dispersive NIWs, which produce a negative frequency shift as described by (4.2). We explore this hypothesis in more detail below. It is useful to note here that the winds excite multiple NIW modes, and so while we hypothesize that these negative shifts come from strongly dispersive waves, we also expect weakly dispersive waves to be present too. In other regions, the frequency shift is weakly positive.

One potential concern is that the observed frequency shifts may be influenced by a vorticity sampling bias inherent to the drifter data. Drifters preferentially sample regions of convergent flow, which are often associated with cyclonic structures (Middleton and Garrett, 1986; Elipot et al., 2010). If weakly dispersive NIWs are present, ray-tracing predicts a positive frequency shift in such regions, potentially introducing a net positive bias. However, we will show below that the spatial patterns in the frequency shift persist even when this bias is taken into account.

The width σ_I of the NIW peak is remarkably uniform across most of the ocean but is elevated in western boundary currents and the ACC (Fig. 4.4b). This suggests an influence of mesoscale eddies, but strongly dispersive waves should be narrowly peaked and the ray-tracing framework for weakly dispersive waves does not provide predictions for spectral width. In the following sections, we use the YBJ model to examine the physical mechanisms responsible for the increased spectral width in high-energy regions.

The NIW kinetic energy is elevated beneath the mid-latitude storm tracks in the North Pacific, North Atlantic, and Southern Ocean, with a notable maximum in NIW kinetic energy in the North Pacific (Fig. 4.4c). This pattern broadly reflects the large-scale distribution of wind energy input into the near-inertial band (Flexas et al., 2019; Alford, 2020; von Storch and Lüschow, 2023). Models suggest some asymmetry in wind work magnitude between the North Pacific and North Atlantic, although the dynamics governing NIW propagation to depth likely also influence the kinetic energy generated by a given wind forcing. We explore these dynamical differences further below.

The patterns of the NIW frequency shift ω_I and peak width σ_I show structure that is spa-

tially aligned with regions of energetic mesoscale turbulence, suggesting that interactions with mesoscale eddies influence the NIW spectral properties. To investigate this influence, we characterize the mesoscale eddy field using satellite altimetry measurements of sea surface height (SSH). We use the delayed-time (DT) 2018 release of the Data Unification and Altimeter Combination System (DUACS) product (Taburet et al., 2019). The mesoscale streamfunction ψ is defined as $\psi = gh/f$, where g is the gravitational acceleration and h is the SSH. We then compute the mesoscale vorticity $\zeta = \nabla^2\psi$ and interpolate it onto all drifter trajectories to obtain concurrent estimates of NIW velocity and mesoscale vorticity. We note that the DUACS product may have inaccuracies near the coast, but drifters in coastal regions make up a very small fraction of the total dataset.

The influence of mesoscale eddies on NIW spectra can be assessed by evaluating NIW spectral properties as a function of ζ. Within each spatial bin, we subdivide the spectrograms based on the mean vorticity along each trajectory, calculated using the same Lanczos window as above. For each vorticity bin, we compute the spectral estimate and extract the spectral fit parameters defined in (4.3).

The spectral properties of NIWs, when averaged globally, show strong evidence of modulation by mesoscale eddies (Fig. 4.5). To quantify this modulation, we perform a linear regression (weighted by the uncertainty on each data point) of ω_I and K to ζ:

$$\frac{\omega_I(\zeta)}{f} = a + b\frac{\zeta}{f}, \tag{4.5}$$

$$K(\zeta) = c + d\frac{\zeta}{f}. \tag{4.6}$$

We find that the frequency shift (Fig. 4.5a) varies with ζ with a slope of 0.39 (95% CI: 0.36 to 0.42) and an intercept of 0.000 (95% CI: -0.001 to 0.001). This slope is in excellent agreement with the values reported by Elipot et al. (2010), although our estimated intercept is lower. The dependence of the spectral width on vorticity (Fig.4.5b) also resembles the findings of Elipot et al. (2010); however, large uncertainties in high-vorticity bins make it difficult to determine whether there is a robust relationship. Globally averaged, we observe a clear concentration of NIW kinetic energy into anticyclonic regions (Fig. 4.5c). The slope of the linear fit is 0.0011 $m^2\ s^{-2}$ (95% CI: 0.0009 to 0.0012 $m^2\ s^{-2}$), and the intercept is 0.0043 $m^2\ s^{-2}$ (95% CI: 0.0042 to 0.0044 $m^2\ s^{-2}$). While a linear fit is not a good model for $K(\zeta)$ at high vorticity, the slope provides a useful summary metric for quantifying the preferential concentration into anticyclones.

Caution is warranted when interpreting global averages of spectral properties. For example, the spectral width is elevated in western boundary currents and the Antarctic Circumpolar Current (ACC), where vorticity is also strong (Fig. 4.4b). The upward curvature in the width–vorticity relationship may therefore reflect a coincidental spatial correlation between broader spectra and larger vorticity, rather than arising from underlying dynamical

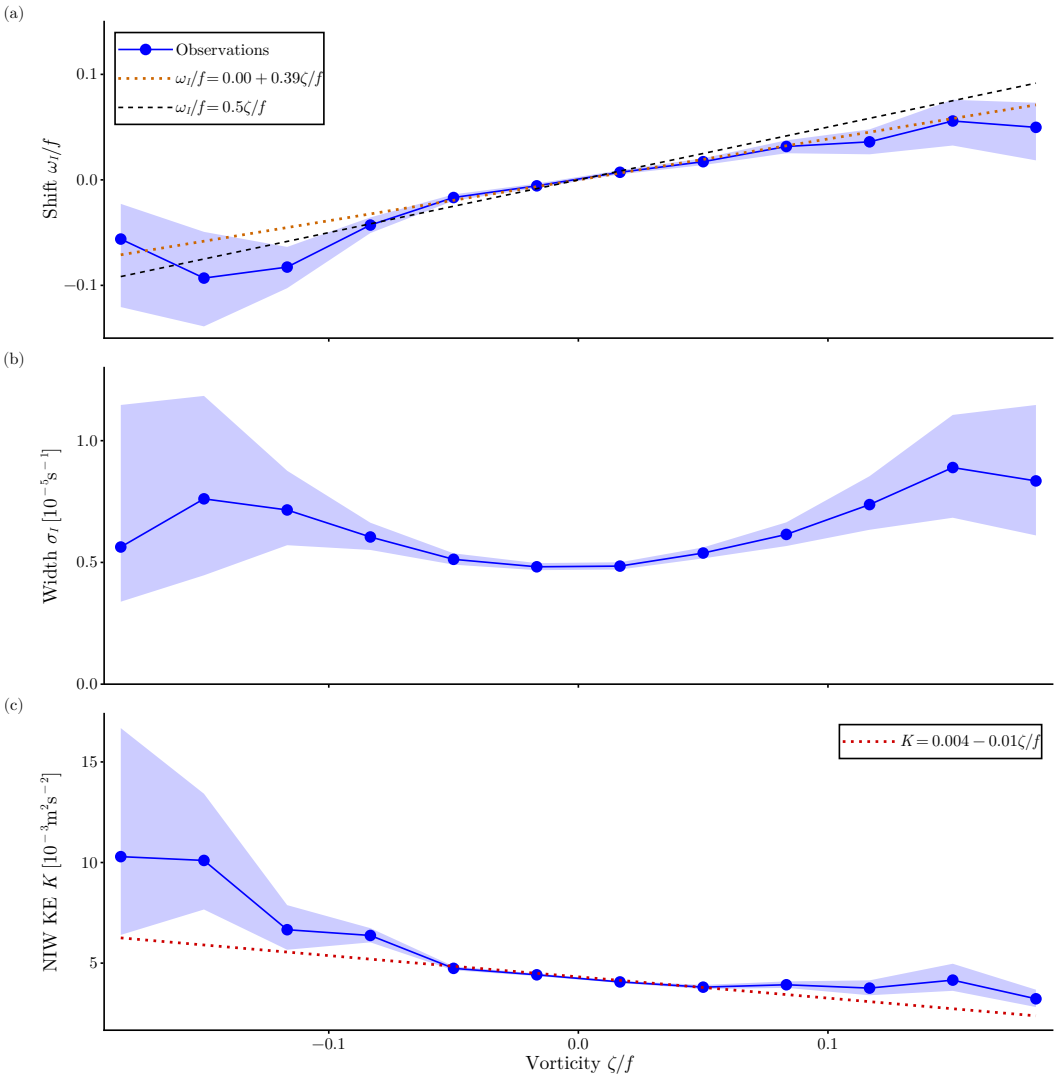

Figure 4.5: Globally averaged dependence of NIW spectral properties on mesoscale vorticity ζ: (a) frequency shift ω_I/f (blue), with the associated linear fit (orange) and the $\zeta/2$ prediction (black), (b) spectral width σ_I (blue), (c) NIW kinetic energy K (blue). Shading indicates 95% confidence intervals.

processes. Similarly, the enhancement of NIW kinetic energy in regions of strong vorticity may influence the global average, although comparisons between strong cyclonic and anticyclonic vorticity remain robust. More generally, interpreting global averages requires care, as they conflate diverse dynamical regimes.

Furthermore, we note that wind forcing typically excites a range of vertical modes (e.g., Gill, 1984; Thomas et al., 2024a), each with different degrees of wave dispersiveness (Fig. 4.1). The drifters feel the combination of these modes and may exhibit, even in a given region, behaviour that is a mix of strongly dispersive low modes and weakly dispersive higher modes. The spectra estimated from the drifters may reflect that mix. The strength of the imprint of strongly vs. weakly dispersive waves depends on their relative amplitudes, and a given diagnostic may be more sensitive to strongly or weakly dispersive waves, requiring nuance in the interpretation.

With these cautionary notes in mind, we calculate linear fits for ω_I and K across the globe. To improve the statistics, we compute $\omega_I(\zeta)$ within each spatial bin and then average these curves to obtain a uniform 5°×5° discretization. A linear fit is then applied to the mean $\omega_I(\zeta)$ curve within each 5°×5° bin. The resulting intercept a exhibits significant spatial structure (Fig. 4.6a). Because of the lack of a strong preference for cyclonic vs. anticyclonic vorticity at the mesoscale (Chelton et al., 2011), the intercept a can be interpreted as the mean frequency shift with any bias in drifter locations toward cyclones removed. Notably, the spatial structure of a (Fig. 4.6a) closely resembles that of the mean ω_I (Fig. 4.4a), suggesting that these patterns are not an artifact of sampling bias. Strong negative intercepts are observed in regions of high eddy kinetic energy, such as western boundary currents and the Antarctic Circumpolar Current. As shown in (4.2), strongly dispersive NIWs are expected to exhibit negative frequency shifts proportional to mesoscale kinetic energy. These negative intercepts may therefore signal the excitation of strongly dispersive NIWs. Positive frequency shifts, on the other hand, are likely due to equatorward propagation of NIWs, although we show below that weakly dispersive waves interacting with mesoscale eddies are also expected to produce positive shifts.

Negative frequency shifts associated with strongly dispersive waves may explain the negative values of the intercept a observed in western boundary currents in the drifter data. A fraction of the near-inertial wind work excites strongly dispersive low modes (e.g., Thomas et al., 2024a), which are associated with a negative frequency shift proportional to the local mesoscale kinetic energy. Whether the full NIW signal also exhibits a negative frequency shift depends on the extent to which the wind excites these modes and the magnitude of their individual frequency shifts. The close spatial alignment between regions of high mesoscale kinetic energy and negative a values leads us to hypothesize that large mesoscale kinetic energy is driving large negative frequency shifts in strongly dispersive waves, which are then expressed in the full signal. To test this idea, we compare the magnitude of the neg-

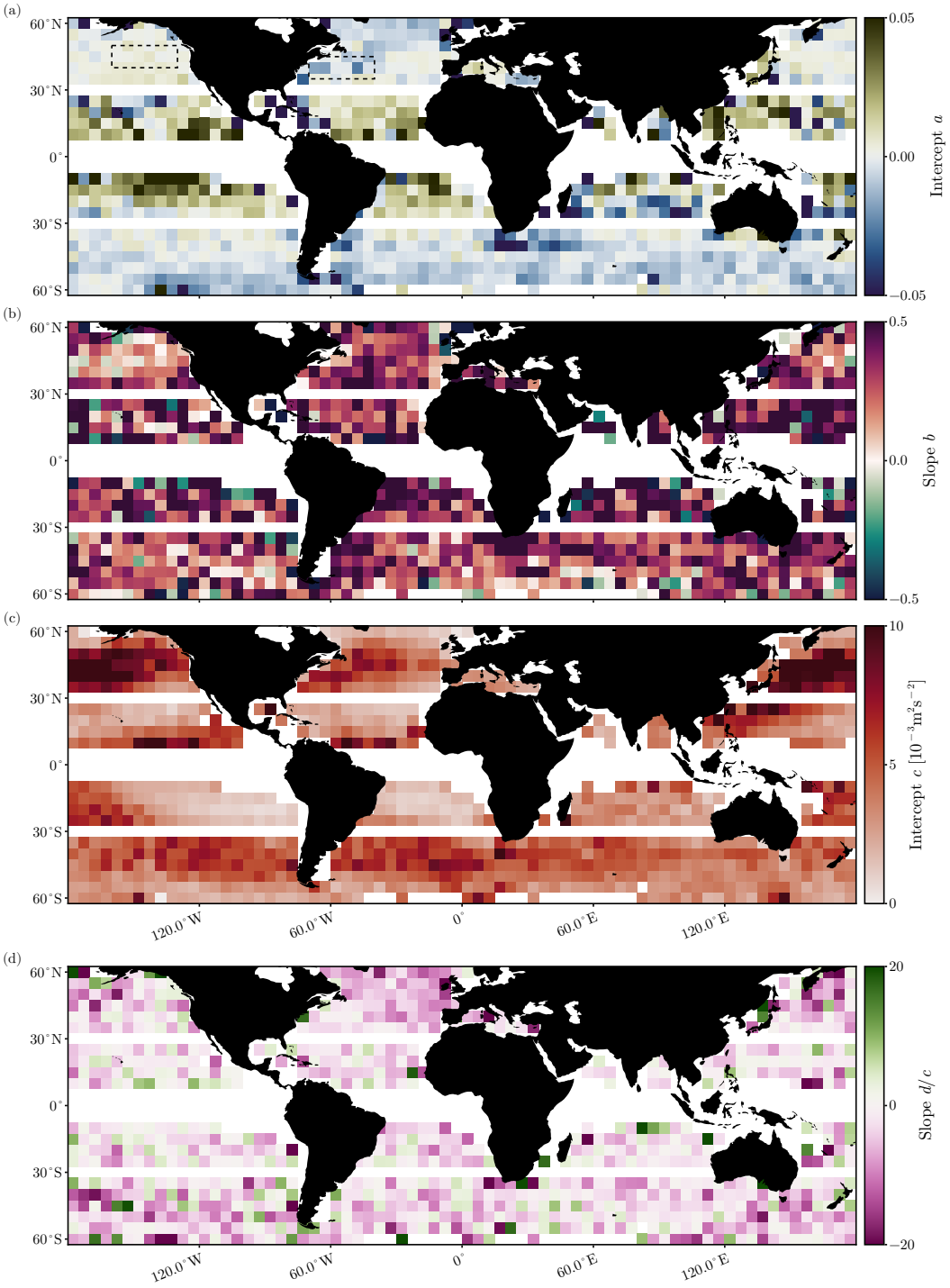

Figure 4.6: Geography of the vorticity dependence of NIW spectral properties. (a) Intercept and (b) slope of the linear fit to the NIW frequency shift ω_I/f as a function of vorticity ζ. The black boxes in (a) outline the two regions to be investigated further in Fig. 4.8. (c) Intercept and (d) slope of the linear fit to the NIW kinetic energy K as a function of vorticity ζ. Note that we show the slope d normalized by the intercept c because otherwise it primarily reflects the patterns seen in panel (c). 95% confidence intervals are shown in Suppl. Fig. 2.

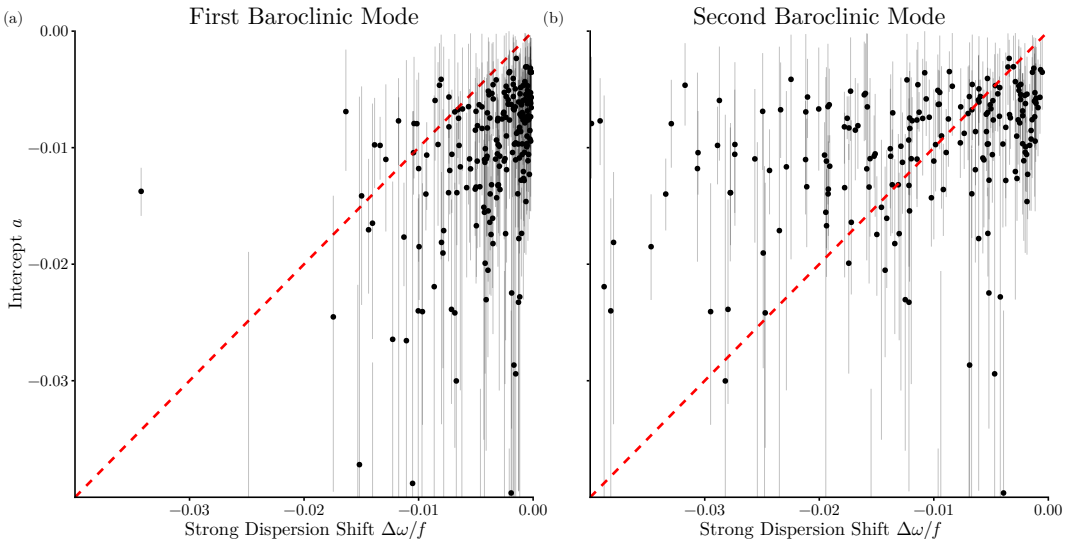

Figure 4.7: Comparison to the strong-dispersion prediction (4.2). Scatter plot of the intercept a of the frequency–vorticity fit in each spatial bin vs. the predicted frequency shift in the strong-dispersion limit for (a) the first and (b) the second baroclinic modes. The 1:1 line is shown in red. The grey bars indicate the 95% confidence interval.

ative intercepts with the prediction from (4.2) for strongly dispersive waves. To estimate this prediction from observations, we require both the mesoscale kinetic energy and the deformation radius. Using the streamfunction from altimetry, we calculate the mesoscale kinetic energy as $\frac{1}{2}|\nabla\psi|^2$. The deformation radius is obtained by solving the baroclinic eigenvalue equation (Smith, 2007) using climatological data from the Estimating the Circulation and Climate of the Ocean (ECCO) state estimate version 4 release 4 (Fukumori et al., 2020; Forget et al., 2015). We perform this calculation for the first two baroclinic modes, which are generally strongly dispersive (Fig. 4.1). The measured a from the drifter data reflects a mixture of frequency shifts from all baroclinic modes, as well as possible contributions from the β-effect, so a one-to-one correspondence is not expected. Nonetheless, we can assess whether the magnitude of the predicted shifts for strongly dispersive waves are consistent with the observed a. The observed a generally lies between the values predicted for the first and second baroclinic modes (Fig. 4.7). It is therefore plausible that the negative values of a result from the presence of strongly dispersive waves.

The slope b, while somewhat noisy, exhibits little spatial variability (Fig. 4.6b). With few exceptions, its value remains close to the global mean. The drifter data thus provides evidence that mesoscale eddies modulate NIW spectra across most of the ocean. The Northeast Pacific stands out as the only region where the slope b is consistently weaker. This region was the site of the Ocean Storms Experiment, whose results similarly indicated weak mesoscale modulation—consistent with our findings. In Section 4.4, we show that, theoretically, the slope is expected to asymptote to a constant value in the weak-dispersion regime, which helps explain the limited spatial structure in b.

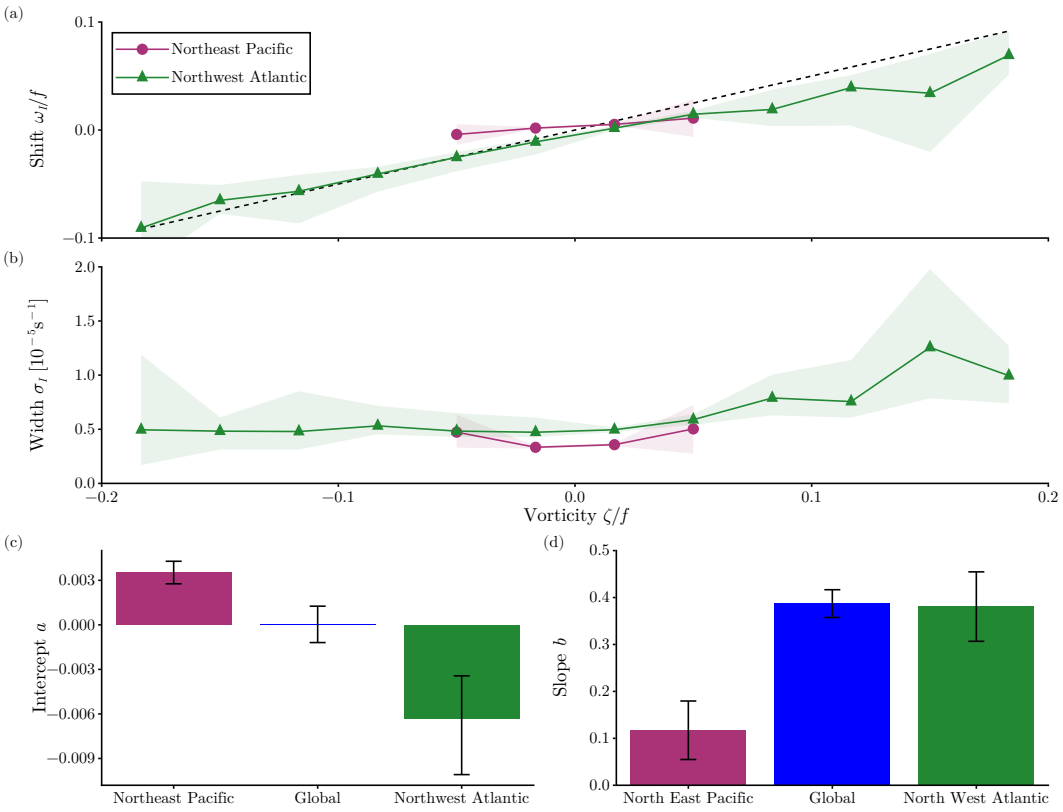

Figure 4.8: Regional vorticity dependence of the NIW spectral properties. (a) NIW frequency shift ω_I/f as a function of vorticity ζ when averaged over all spectra in the Northeast Pacific (purple) and the Northwest Atlantic (green) regions marked in Fig. 4.6a. The dashed black line shows the ray-tracing $\zeta/2$ prediction. (b) Same but for the NIW spectral width σ_I/f. (c) Intercept of the linear fits in (a) and the global average (blue) for comparison. (d) Same but for the slope of the linear fit.

The intercept c (Fig. 4.6c) closely mirrors the spatial distribution of NIW kinetic energy in the ocean (cf., Fig. 4.4c), and the associated NIW wind work. The slope d, while somewhat noisy, is predominantly negative (Fig. 4.6d), indicating that the concentration of NIWs into anticyclones is a widespread feature of the global ocean. That said, there are notable regions where the slope d weakens. In particular, the Gulf Stream region exhibits a very modestly negative slope, indicating weakened concentration into anticyclones, despite the presence of strong eddies. In regions with strong eddies, which have relatively weaker wave dispersion (Fig. 4.1), the tendency of waves to concentrate in anticyclones may be disrupted by advective straining (Rocha et al., 2018), or regions of wave concentration may be shifted from the anticyclones by baroclinicity in the mean flow (Whitt and Thomas, 2013). While this map shows that anticyclonic concentration is a common aspect of NIW dynamics, it also highlights that the effect depends on wave dispersiveness in complex ways.

As the slope b is somewhat noisy, and to better understand the behavior of the spectral width, we average the spectra over two representative regional domains (shown as black boxes in Fig. 4.6a). These two regions differ markedly in the characteristics of their me-

soscale eddy fields, and are chosen to highlight characteristics of the two different wave regimes. In the Northeast Pacific box, eddies are weak, while the Northwest Atlantic box encompasses the energetic Gulf Stream rings. In the Northeast Pacific, we observe smaller slopes and a positive frequency shift; in the Northwest Atlantic, we find steeper slopes and a negative frequency shift. The regional averages support this distinction (Fig. 4.8a). The spectral width is greater in the Northwest Atlantic than in the Northeast Pacific (Fig. 4.8b). The weak slope observed in the Northeast Pacific appears to be uncommon across the global ocean, with the possible exception of a few regions in the Southern Ocean, although the statistical uncertainty is higher there. The weak eddy field in the Northeast Pacific corresponds to higher wave dispersiveness overall, favoring strongly dispersive waves and reducing mesoscale modulation. In the Northwest Atlantic, we see the negative frequency shift indicative of strongly dispersive waves, but also clear modulation of the frequency shift by mesoscale vorticity, consistent with weakly dispersive wave behavior. Once again, we observe a complex picture in which the winds excite multiple vertical modes, each with a different degree of dispersiveness. The spectral characteristics reflect the aggregate influence of all these modes.

We have identified several characteristic properties of NIW spectra on a global scale that appear linked to interactions with mesoscale eddies. Regions of negative frequency shift are collocated with energetic western boundary currents and the ACC. Additionally, the slope of the frequency shift with vorticity shows remarkably little variation across the ocean, typically remaining close to the global average of $\sim$0.4. In a few localized regions, the slope flattens to near zero. Finally, the spectral width is elevated in energetic regions. To more quantitatively investigate the physical origins of these patterns, we examine the spectral properties of NIWs in an idealized model of NIW–mesoscale interactions.

4.4 Idealized Simulations of NIWs

Simulations of the YBJ equation provide an idealized framework for investigating the impact of mesoscale interactions on NIW spectral properties across different regimes, and for assessing the extent to which NIW–mesoscale interactions shape the observed NIW spectra from drifters. Xie and Vanneste (2015, hereafter XV) proposed a model that couples the YBJ equation to a quasi-geostrophic (QG) flow. We simulate the XV model across a range of wave dispersiveness values. Using these simulations, we advect Lagrangian particles through the NIW field to generate synthetic drifter observations. These synthetic trajectories can then be used to calculate NIW spectral properties under controlled conditions.

Following Rocha et al. (2018), the XV model for a given vertical mode is:

$$\frac{\partial \phi}{\partial t} + \mathrm{J}(\psi, \phi) + \frac{i\zeta}{2}\phi - \frac{i f \lambda^2}{2}\nabla^2\phi = -\nu_\phi \nabla^4 \phi, \tag{4.7}$$

$$\frac{\partial q}{\partial t} + \mathrm{J}(\psi, q) = -\kappa_q \nabla^4 q, \tag{4.8}$$

$$q = \nabla^2\psi + q_w, \quad q_w = \frac{1}{f_0}\left[\frac{1}{4}\nabla^2|\phi|^2 + \frac{i}{2}\mathrm{J}(\phi^*, \phi)\right], \tag{4.9}$$

where q is the potential vorticity (PV), $\nabla^4 = \nabla^2\nabla^2$ is the biharmonic operator, ν_ϕ is the hyper-diffusivity for NIWs, and κ_q is the hyper-diffusivity for the QG flow. The biharmonic terms are included for numerical stability. This formulation represents a 2D version of the XV model in which the mesoscale flow is assumed to be barotropic, and the NIW velocity has been expanded in vertical baroclinic modes. The term q_w represents the effects of the back reaction that the waves have on the mean flow; it can be thought of a wave-induced potential vorticity. These equations are solved on a doubly periodic domain using the Dedalus pseudospectral solver (Burns et al., 2020). We also note that these are spin-down, unforced turbulence simulations. The parameters used are provided in Appendix B. We vary ε^2 in the simulations by varying the deformation radius λ, given the same f and streamfunction magnitude Ψ.

We begin by evolving the QG flow for 25 days without waves (i.e., $\phi = 0$). After this spin-up period, we introduce a horizontally uniform NIW field with magnitude U_w. The waves are then allowed to evolve for an additional 60 days. We repeat the simulations for various values of wave dispersiveness. The qualitative structure of the wave field varies significantly with dispersiveness. For $\varepsilon \gg 1$, the waves exhibit large-scale structure resembling the streamfunction (Fig. 4.9a). For $\varepsilon \ll 1$, in contrast, the waves take on smaller-scale features more akin to the vorticity (Fig. 4.9b).

For each experiment, we advect Lagrangian particles to mimic the drifters. Particles are seeded in the model domain with a spacing of 25 km and advected forward and backward for 10 days, yielding a total 20-day trajectory per particle. New particles are reseeded every 10 days. We use the Ocean Parcels package (Delandmeter and Van Sebille, 2019) to perform the Lagrangian advection. The advection velocity is derived from ψ, which represents the Lagrangian-mean streamfunction that transports tracers (Wagner and Young, 2015). Particles sample the NIW velocity ϕe^{-ift} at hourly intervals.

The dependence of the frequency shift and spectral width on vorticity varies markedly with ε. In the strong-dispersion regime ($\varepsilon \gg 1$), there is no modulation of either the shift or the width by vorticity. In contrast, in the weak-dispersion regime ($\varepsilon \ll 1$), both quantities show clear modulation with vorticity (Fig. 4.10a,b): the frequency shift exhibits a positive slope with vorticity, while the width is enhanced in cyclonic regions. The intercept a is negative in the strong-dispersion regime and variable but generally positive in the weak-dispersion

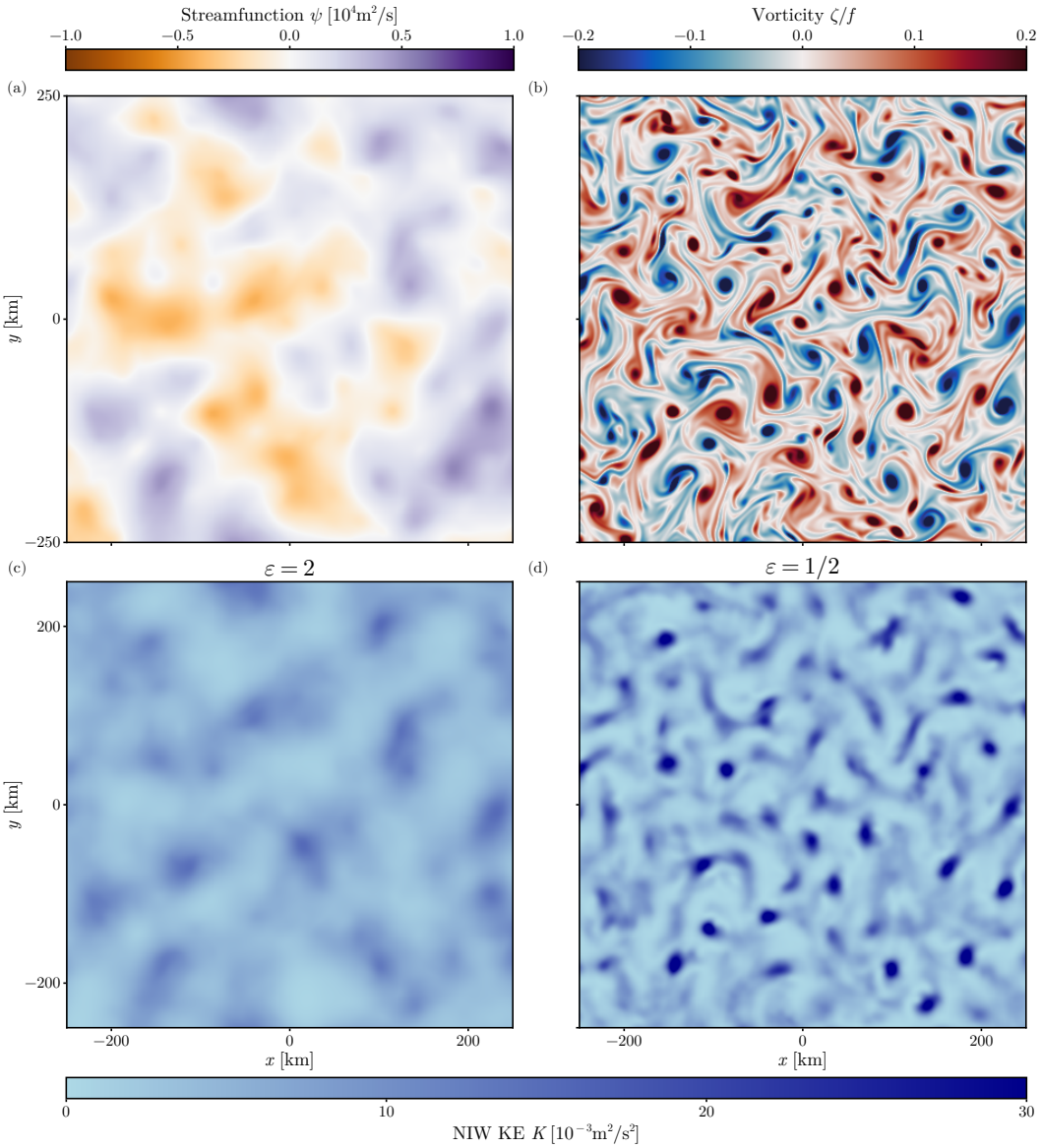

Figure 4.9: Idealized XV simulations of NIW–mesoscale interaction. (a) Snapshot of the streamfunction ψ in the simulations with $\varepsilon = 2$ at $t = 4$ days. (b) Snapshot of the corresponding vorticity. (c) Snapshot of the corresponding NIW kinetic energy density. Note that the structure resembles the streamfunction in panel (a). (d) Same but for $\varepsilon = 1/2$. Note that there is more small-scale structure that resembles the vorticity field. (At this early stage, the vorticity looks the same for $\varepsilon = 2$ and $1/2$. The impact of q_w on the QG flow has not yet had a chance to cause the two simulations to drift apart. As such, we only show the vorticity for $\varepsilon = 2$.)

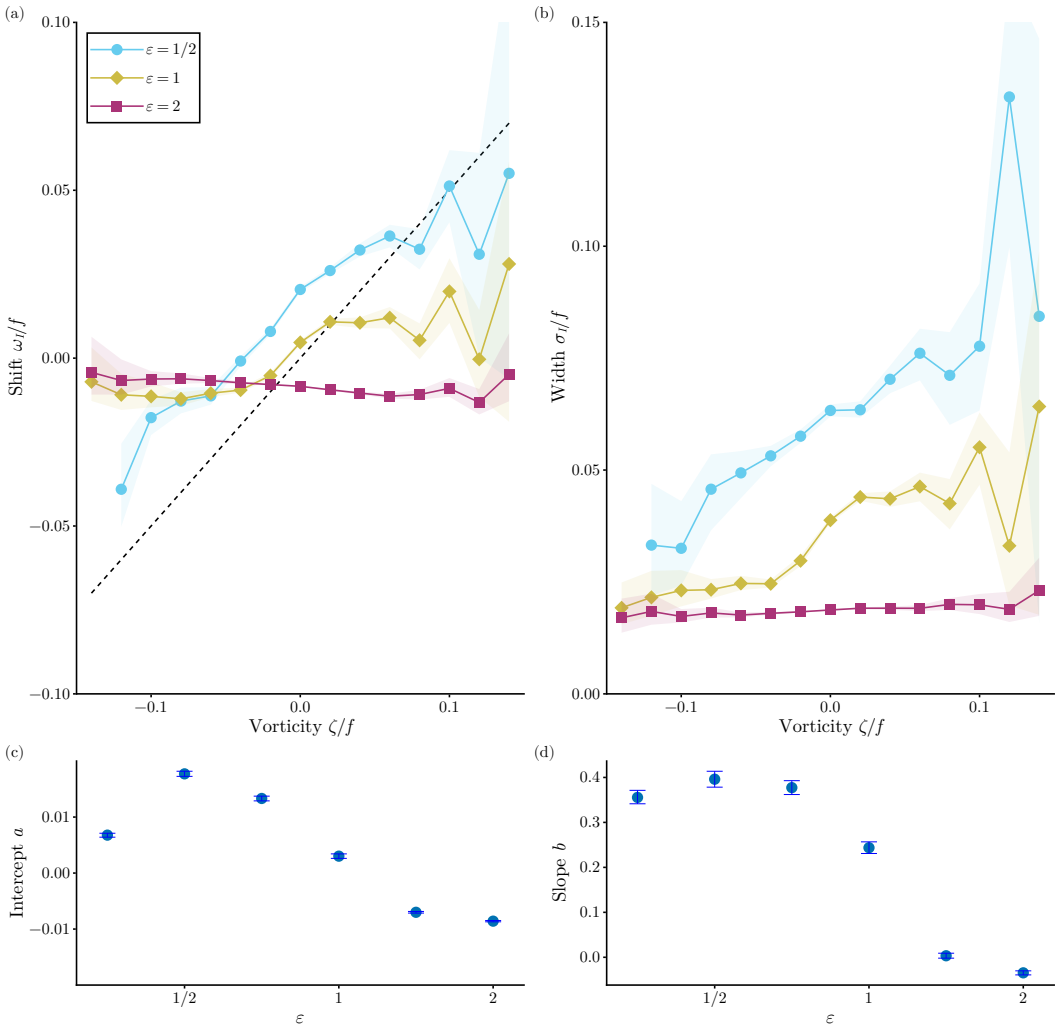

Figure 4.10: Spectral characteristics of the NIW peak in the XV simulations. (a) NIW frequency shift ω_I/f as a function of vorticity ζ for $\varepsilon = 1/2$ (weak dispersion, blue), $\varepsilon = 1$ (transition case, yellow), and $\varepsilon = 2$ (strong dispersion, purple). The $\zeta/2$ line is shown in black. (b) Same but for the NIW spectral width σ_I. (c) Intercept a of the linear fit as a function of ε. (d) Same but for the slope b.

regime (Fig. 4.10c). The slope approaches zero in the strong-dispersion limit and asymptotes to a value of ∼0.4 in the weak-dispersion limit (Fig. 4.10d).

These XV simulations produce spectra that exhibit many of the salient features observed in the NIW spectra from drifters, supporting the idea that these features arise from NIW–mesoscale interactions. While idealized, however, the simulations are still complicated, making it difficult to isolate the physics responsible for the imprint of mesoscale eddies on NIW spectra. To better illuminate these dynamics, we turn to a much simpler example: NIWs interacting with a single vortex dipole.

4.5 NIWs in an Idealized Vortex Dipole

The goal of this section is to understand the physics of NIW–mesoscale interactions and their effect on NIW spectra in the simplest possible context. Conn et al. (2025) adopted a spectral approach to solving the YBJ equation, calculating the eigenvalues and eigenvectors of the YBJ operator—a useful method for studying NIW spectra. We begin by non-dimensionalising the YBJ equation and rewriting it as

$$i\frac{\partial\tilde{\phi}}{\partial\tilde{t}} = H\tilde{\phi} = -\frac{\varepsilon^2}{2}\nabla^2\tilde{\phi} - i\,\mathrm{J}(\tilde{\psi},\tilde{\phi}) + \frac{\tilde{\zeta}}{2}\tilde{\phi}, \tag{4.10}$$

where H is the YBJ operator, the tildes indicate non-dimensionalized fields, and the operators are understood to take derivatives with respect to the non-dimensional variables. We then consider the spectrum of the operator H by solving for the eigenvalues and eigenfunctions of H,

$$H\tilde{\phi}_n = \tilde{\omega}_n\tilde{\phi}_n, \tag{4.11}$$

where $\tilde{\phi}_n$ is an eigenfunction, $\tilde{\omega}_n$ is the associated eigenvalue, and n is a label for the eigenfunction. Throughout this section, we consider a simple vortex dipole (Fig. 4.11a Asselin et al., 2020; Conn et al., 2025),

$$\tilde{\psi} = \frac{1}{2}\left(\sin\tilde{x} - \sin\tilde{y}\right), \tag{4.12}$$

consisting of a cyclone and anticyclone on a doubly periodic domain with side length 2π.

In the strong-dispersion limit $\varepsilon \gg 1$, the nearly uniform mode dominates and results in a strongly peaked spectrum at a weakly shifted frequency as described by (4.2). Therefore, we here focus our attention on the weak-dispersion limit. For the moment, let us ignore advection and write the eigenvalue equation as

$$-\frac{\varepsilon^2}{2}\tilde{\nabla}^2\tilde{\phi}_n = \left(\tilde{\omega}_n - \frac{\tilde{\zeta}}{2}\right)\tilde{\phi}_n. \tag{4.13}$$

From this equation, one can see that $\tilde{\zeta}$ acts like a potential in the Schrödinger equation, and for a given mode with eigenvalue $\tilde{\omega}_n$, the eigenfunction is oscillatory where $\tilde{\omega}_n > \tilde{\zeta}/2$ and evanescent where $\tilde{\omega}_n < \tilde{\zeta}/2$. In other words, lower-frequency modes are screened out in

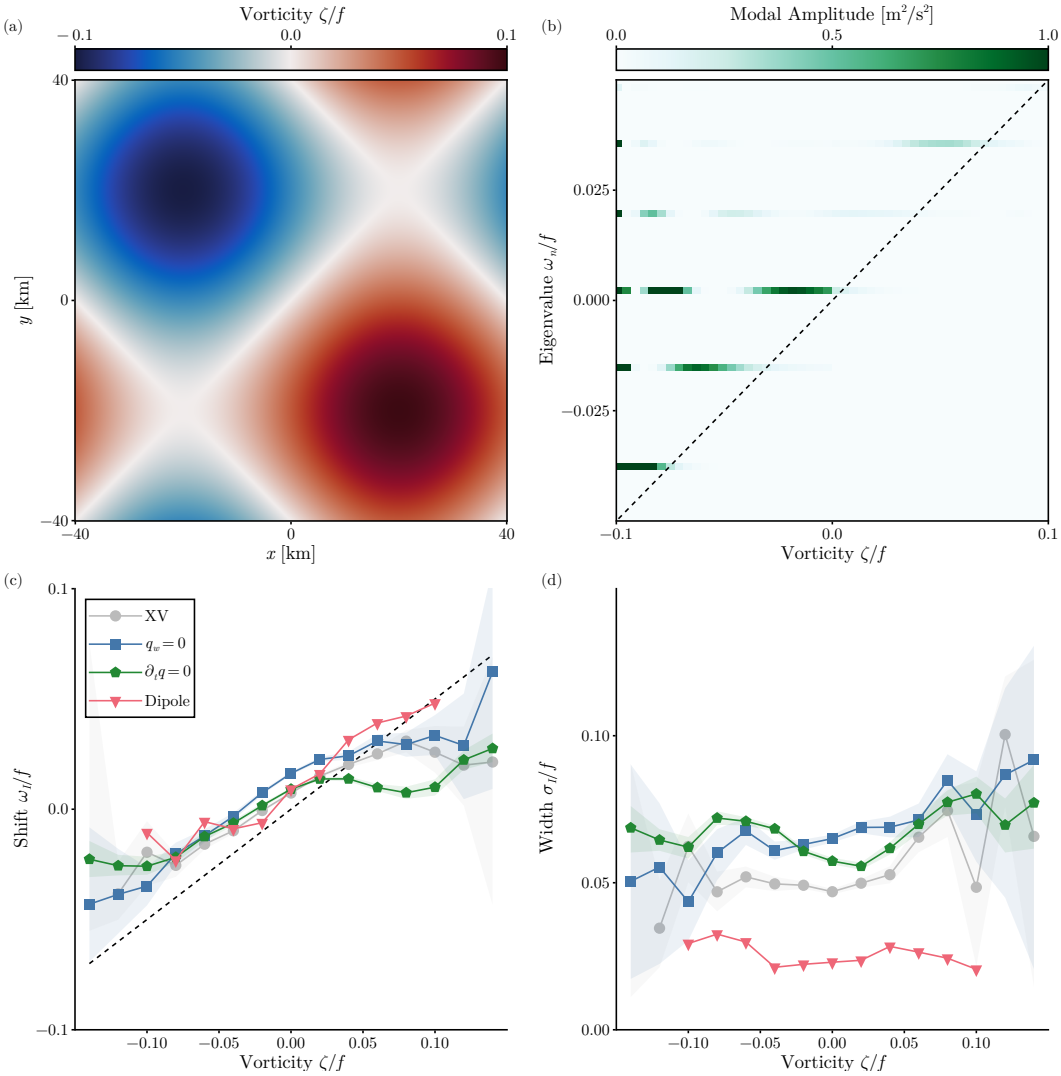

Figure 4.11: NIW spectral characteristics for progressively idealized simulations. (a) Vorticity of the dipole flow. (b) Modal amplitude as a function of vorticity for each eigenmode of the dipole flow. (c) NIW frequency shift ω_I as a function of ζ for the simulations with the XV coupling turned off (blue), with the streamfunction evolution turned off (green), and for the dipole flow (pink). The $\zeta/2$ line is shown in black. (d) Same but for the NIW spectral width σ_I. Note that we have re-dimensionalized the dipole flow with a streamfunction magnitude Ψ that matches the XV simulations and a domain length of 80 km. Furthermore, the dipole lines come from YBJ simulations with the dipole as the background field. We also calculated these lines from the eigenmodes described above and found that the two methods agree.

regions of higher vorticity. At a given vorticity $\tilde{\zeta}$, one should thus see only modes that have a frequency of $\tilde{\zeta}/2$ or greater. While a number of modes contribute to the NIW frequency spectrum at low vorticities, a large-scale wind forcing, here modeled as a uniform forcing, projects most strongly onto the low modes. As a result, the NIW frequency increases monotonically with $\tilde{\zeta}$, with a slope of the frequency–vorticity curve close to $1/2$. This argument is admittedly hand-wavy and neglects advection, but the more rigorous treatment in Conn et al. (2025) supports this intuition.

To give a concrete example—now including advection—we compute the eigenfunctions and eigenvalues for the dipole flow. The numerical procedure is described in Conn et al. (2025). We also compute the projection coefficient of a uniform initial condition onto each mode, and we define the modal amplitude as the mean, taken along contours of constant vorticity, of the modulus squared of each eigenmode multiplied by its projection coefficient. The results clearly illustrate the screening of modes with $\tilde{\omega}_n$ less than the local $\tilde{\zeta}/2$ (Fig. 4.11b). The strong projection onto the lowest eigenmodes ensures that the mean frequency remains close to the $\tilde{\zeta}/2$ line, even in the core of the anticyclone (Fig. 4.11c).

The dipole calculation therefore provides a physical explanation for the observed modulation of the frequency shift by vorticity. This modulation occurs across much of the ocean (Fig. 4.6b), implying the near-universal presence of weakly dispersive wave modes that are refracted by mesoscale vorticity. The dipole calculation also predicts, however, that the spectral width should be elevated in anticyclones (Fig. 4.11b,d), whereas the simulations and, to a lesser extent, the observations show the opposite. To investigate this discrepancy, we return to the XV model from Section 4.4, running it with varying levels of complexity.

First, we disable the influence of NIWs on the mesoscale by setting $q_w = 0$ in the XV equations. The QG flow continues to evolve, but there is no feedback from the waves. Next, we run a simulation in which we turn off the time evolution of the QG flow altogether: the initial condition is the same as before, but the flow is held fixed in time after the spin-up. Finally, we perform a simulation in which we replace the initial condition with the dipole flow (redimensionalized to have $\mathrm{Ro} = 0.1$). This is a stationary solution to the QG equations and therefore does not evolve in time. These three simulations form a hierarchy of complexity, ranging from the simplest dipole case to the full XV model.

Setting $q_w = 0$ has minimal impact on the results compared to the full XV simulation, leading only to a slight increase in the spectral width (Fig. 4.11c,d). The $\partial_t q$ simulation also leaves the frequency–vorticity relation largely unchanged, except that there is a drop in the frequency shift at positive ζ (Fig. 4.11c). It shows similar spectral width as the XV and $q_w = 0$ simulations. The dipole simulation recovers a frequency–vorticity relation with a slope close to $1/2$, but it has a markedly decreased spectral width and no increase with ζ (Fig. 4.11d). Instead, as expected from the modal picture, the spectral width is greatest at negative ζ.

It appears that the key complexity required to produce spectra that qualitatively resemble the drifter observations is the presence of a sea of eddies of varying amplitude. Even when the eddies are stationary—just one step up in complexity from the dipole simulations—we observe behavior in both the frequency shift and the spectral width that resembles the observations. In the dipole flow, averaging over vorticity amounts to averaging within a single eddy, whereas in the turbulent QG flow, the averaging spans multiple eddies with varying structures. In a sea of eddies, a given cyclone may have stronger cyclones nearby, meaning it is no longer necessarily at the top of the $\zeta/2$ landscape that acts like a potential. Eigenmodes with frequencies higher than that of the local cyclone can be excited and felt within it, thereby increasing the spectral width. The same does not hold for anticyclones, because modes with energy lower than the local $\zeta/2$ are screened. The result is an asymmetry in spectral width between cyclones and anticyclones as observed in the simulations. The behavior of the spectral width, therefore, is not governed solely by NIWs interacting with isolated eddies, but instead emerges from interactions within a sea of eddies.

4.6 Discussion

The drifter observations, interpreted in the context of NIW–mesoscale interactions, reveal a complex picture of NIW behavior. Spatial variability in the mesoscale flow is clearly imprinted onto the NIW spectral properties. Even within a single region, however, we observe signatures of both weakly and strongly dispersive waves. This arises from the wind forcing exciting multiple vertical modes, each characterized by a different degree of wave dispersiveness. The drifters sample a combination of these modes and their associated behaviors. While it is beyond the scope of this manuscript to determine which vertical modes are excited in each region and how dispersive they are, this remains a key factor in the evolution of NIWs. Thomas et al. (2024a) performed such a projection in regions of the North Atlantic and North Pacific, and Conn et al. (2025) calculated climatological wave dispersiveness for the first four baroclinic modes (see Fig. 4.1). Extending Thomas et al.'s calculations to a global scale would provide valuable insight into how the structure of wind forcing and background stratification influence the spectral characteristics of NIWs.

Elipot et al. (2010) attributed the positive shift in NIW frequency equatorward of 30° latitude to be due to the β-effect, which causes equatorward propagation of NIWs and therefore produces a positive shift. However, we also saw that a positive shift could be generated by weakly dispersive waves (Figures 4.10, 4.11). The global pattern and the increasingly positive shifts toward the equator, where β is larger, support the equatorward propagation mechanism. Garrett (2001) calculated the meridional distance that NIWs propagate in the time it takes them to reflect off the seafloor and return to the surface, finding ~400 km except for very close to the equator and pole. The associated frequency shift, normalized by f, increases markedly toward the equator and is of order 0.1. The observed frequency shift is somewhat smaller but has the right order of magnitude and pattern. The distance is not large

enough for NIWs generated under mid-latitude storm tracks to contribute substantially to subtropical NIW energy, but it is plausible that the subtropical NIW peak consists of a mix of NIWs generated locally and up to a few degrees poleward (cf., e.g., Raja et al., 2022). The origin of these frequency shifts should be further investigated using realistic simulations or in-situ field campaigns, which are better suited than surface drifters to distinguish the sources of NIW energy.

4.6.1 Caveats and Sources of Uncertainty

In this study, we have considered time means of all spectral quantities and have therefore ignored any potential seasonality in the NIW spectral properties. Seasonal variations in mesoscale eddies and the deformation radius are generally weak over much of the ocean (cf., e.g., Chelton et al., 1998) and so these are unlikely to significantly alter these results. However, both the wind forcing and stratification exhibit substantial seasonal variability (cf., e.g., Alford and Whitmont, 2007; Sallée et al., 2021). This seasonal variability could impact the projection of the wind forcing onto the different modes, changing the overall properties of the NIW spectra. A global-scale analysis of the seasonality of NIW spectral characteristics—similar to the approach taken in this paper—would likely be difficult at present due to the limited number of observations across much of the ocean. Regional analyses may prove feasible.

Strictly, the decomposition of NIWs into vertical modes is only valid under the assumption of a barotropic background flow. This assumption is clearly violated in the ocean, where mesoscale eddies often exhibit significant vertical structure. Interestingly, however, accounting for baroclinicity in the mean flow does not appear necessary to explain the spectral characteristics observed by the drifters over much of the ocean, at least qualitatively. There are however some regions where baroclinicity may impact the results. In particular, the lack of strong concentration into anticyclones in the Gulf Stream could arise from the strongly baroclinic eddies in the region. As explained by Whitt and Thomas (2013), the picture of NIW concentration into anticyclones breaks down in the presence of strongly baroclinic eddies. In such a scenario, the locations of minimum NIW frequency can be shifted from locations of minimum vorticity by strong lateral density gradients, which can lead to concentration into the periphery of eddies. More work is needed to understand the spectral properties expected for NIWs interacting with a strongly baroclinic mean flow.

In the real ocean, vorticity is a multiscale field, which we characterize using the heavily smoothed altimetry product. This vorticity field is further averaged along drifter trajectories. One might worry that substantial submesoscale vorticity is being missed and could be influencing the NIWs. We argue that this may be so, but dispersion acts as a natural filter for vorticity. In spectral space, dispersion scales like k^2, where k is the wavenumber. Thus, while vorticity tends to be stronger at smaller scales, so is dispersion. There must exist a scale below which dispersion is sufficient to balance refraction. The fact that we observe such large

slopes in the frequency shift as a function of the mesoscale ζ suggests that this relevant scale is by and large resolved by the altimetry. Indeed, shortening the length of drifter trajectory chunks reduces the slope; for example, using 6 day chunks lowers the slope to $\sim$0.25 (not shown). This implies that, on average, NIWs are not responding to local, small-scale vorticity features. We add the caveat that in regions with a small deformation radius, it is likely that scales relevant to the NIW evolution are not full resolved by altimetry, although data from SWOT may prove useful in this regard (e.g. Rolland et al., 2025).

4.6.2 Other Implications

The observed spectral characteristics of NIWs have implications for filtering NIWs from raw signals that contain multiple types of motion. Rama et al. (2022a) recently proposed using an adaptive filter in which the effective frequency $f+\zeta/2$ serves as a lower bound on a high-pass filter to separate NIWs from other flows. Our results suggest that this approach may not always be appropriate. This expression for the lower frequency limit is derived from ray-tracing, which is valid only in the weakly dispersive limit. Our results show clear evidence of strongly dispersive waves as well, for which there is no comparable modulation of the NIW frequency. Applying such an adaptive filter would introduce a cyclonic–anticyclonic bias when extracting strongly dispersive NIWs. Moreover, even for weakly dispersive waves, the $\zeta/2$ scaling is not universal. The drifter observations typically exhibit a slope closer to $\sim$0.4, with non-negligible NIW variance below the $f+\zeta/2$ cut-off. Using this cut-off would result in severe underestimation of NIW energy in cyclonic regions compared to anticyclonic ones. Small-scale vorticity, as discussed above, further complicates this approach. An adaptive approach that first locates and characterizes the NIW peak seems more promising.

Finally, we note that these results have implications for NIW-induced mixing on a global scale. Numerous observational studies have shown that the concentration of NIWs into anticyclones can lead to enhanced NIW-induced mixing at depth (e.g., Kawaguchi et al., 2016; Martínez-Marrero et al., 2019; Sanford et al., 2021). Our results indicate that this concentration is fairly ubiquitous across the ocean. Therefore, in regions of high NIW energy, the interaction between mesoscale eddies and NIWs may play a significant role in shaping patterns of upper-ocean mixing. Future work should aim to quantify the magnitude of this effect.

4.7 Conclusions

Using observations of near-inertial waves from the global array of drifters, we have mapped the geography of NIW–mesoscale interactions across the world ocean. Our results indicate that mesoscale eddies influence the spectral characteristics of NIWs nearly everywhere. We also found that the YBJ equation, which describes the evolution of NIWs in a background mesoscale eddy field, explains much of the observed phenomenology of NIW spectra. Additionally, we observed evidence of both weakly and strongly dispersive NIWs, suggesting

that wind forcing excites vertical modes for which dispersion has varying importance.

To quantify the modulation of NIWs by mesoscale eddies, we considered a linear fit to the relationship between the NIW frequency shift and the vorticity. The slope of this fit is $\sim$0.4 in the global mean and takes a similar value across much of the ocean—except in the Northeast Pacific and possibly some regions of the Southern Ocean, where the slopes are close to zero. Typical slopes are close to the 0.5 value predicted by ray-tracing, although the XV simulations showed that the slope should asymptote to a somewhat smaller constant value in the weak-dispersion limit, offering an explanation for the lack of spatial variability in the slope on a global scale. Additionally, the dipole flow demonstrated that this slope arises from the screening of eigenmodes of the YBJ equation in regions in which the eigenvalue is less than $\zeta/2$.

In contrast to the slope, the intercept exhibited much greater spatial structure when mapped globally. In highly energetic regions such as the western boundary currents and the ACC, the intercept was negative, while elsewhere it was weakly positive. This close spatial collocation led us to consider strongly dispersive waves as a possible explanation for the negative frequency shifts. Theoretical predictions for the frequency shift in the strong-dispersion limit were found to be of similar magnitude as the observed shifts, acknowledging that a one-to-one correspondence is not expected. The pattern of positive shifts pointed to the β-effect as a likely source, although weakly dispersive waves are also expected to produce positive shifts.

The NIW spectral width also showed a striking correspondence with the highly energetic western boundary currents and the ACC, where the width was elevated. This can be understood by considering the eigenfunctions of the YBJ equation. In a stronger eddy field, the potential wells are deeper, allowing more modes to be excited and thereby increasing the spectral width. The width typically increases with vorticity, a behavior that is expected for a turbulent field of eddies of varying size and reproduced by idealized XV and YBJ simulations.

Finally, we also considered how NIW kinetic energy depends on vorticity. The global mean showed a clear concentration into anticyclones, while the global distribution of mean NIW kinetic energy reflected the patterns of large-scale forcing. Examining the slope of NIW kinetic energy versus vorticity across the globe, we found that anticyclonic concentration is typical throughout much of the ocean. However, a few notable regions—the Gulf Stream region, in particular—exhibited a weak dependence of kinetic energy on vorticity, despite being highly energetic. This suggests that while concentration into anticyclones is widespread, strong advection or baroclinicity can disrupt this tendency.

Parameter	Value
Domain length L	$4\pi \times 10^5$m
Number of modes N	1024
Coriolis frequency f	10^{-4} s^{-1}
NIW speed U_w	0.1 m s^{-1}
Eddy speed U_e	0.05 m s^{-1}
Eddy wavenumber k_e	$16\pi \times 10^{-6}$m^{-1}
Biharmonic viscosity ν_w	5×10^6 m^4 s^{-1}
Biharmonic diffusivity κ_e	5×10^6 m^4 s^{-1}
Timestep Δt	25 s
Simulation run time T	5.184×10^6 s

Table 4.1: Parameters for the XV simulation

4.8 Appendix A: Least-Squares Fit

Given some spectral estimate $\hat{S}$, we wish to find the parameters in Equation 4.3 that minimize the residuals between $\hat{S}$ and the model. Writing the model as $S(\omega, p)$, where p represents the set of model parameters, the residuals are defined as

$$R(p) = \sum_n \left[\hat{S}_n - S(\omega_n, p)\right]^2, \tag{4.14}$$

where ω_n are the discrete frequencies resolved by the data. We use the Levenberg–Marquardt algorithm to find the set of parameters p^* which minimizes $R(p)$.

4.9 Appendix B: Simulation Parameters

For most of the simulation parameters, we follow Rocha et al. (2018). These can be found in Table 4.1. We also follow Rocha et al. (2018) in initializing the QG flow by specifying the Fourier transform of ψ as

$$\hat{\psi}(\boldsymbol{k}) = \frac{C}{\sqrt{|\boldsymbol{k}|\left[1 + \left(\frac{|\boldsymbol{k}|}{k_e}\right)^4\right]}}, \tag{4.15}$$

where the hat indicates the Fourier transform, $|\boldsymbol{k}|$ is the magnitude of the wavevector, k_e is an eddy wavenumber, and C is a normalization constant. The value of C is chosen such that the domain-average kinetic energy density of the simulation is $U_e^2/2$.

The value of ε is varied by changing the deformation radius λ. We choose λ based on the RMS value of the streamfunction and f to give the desired value of ε according to $\varepsilon^2 = f\lambda^2/\Psi$.

4.10 Supplemental Information: Confidence Intervals

The spectral estimates, generated from multiple spectrograms, follow a χ^2 distribution. The fitted parameters are a complicated non-linear function of the spectral estimate and so we

do not have an analytical form for their probability distributions. However, we can draw random samples from the distribution of the spectral estimate, put these through the fitting procedure, and build empirical distribution functions for the fitted parameters. Here we show the 95% confidence interval half widths for the globally mapped parameters.

When using all spectrograms in a given box, the confidence intervals are generally quite narrow (Suppl. Fig. 1). The exception is near the equator and in coastal regions where the confidence intervals can become large.

When binning the spectrograms by vorticity too, we necessarily have less spectrograms being averaged to get a spectral estimate. As a result, the confidence intervals are wider for properties that depend on the vorticity (Suppl. Fig. 2). In general, the intercepts have a relatively narrow confidence interval, but the interval can be come large for the slopes. These plots reflect the distribution of drifter numbers, with the statistics being much better in the Atlantic where we have more drifter observations. For the frequency slope, the values are quite uncertain equatorward of 30°, and in much of the Southern Ocean. The same is generally true for the KE slope.

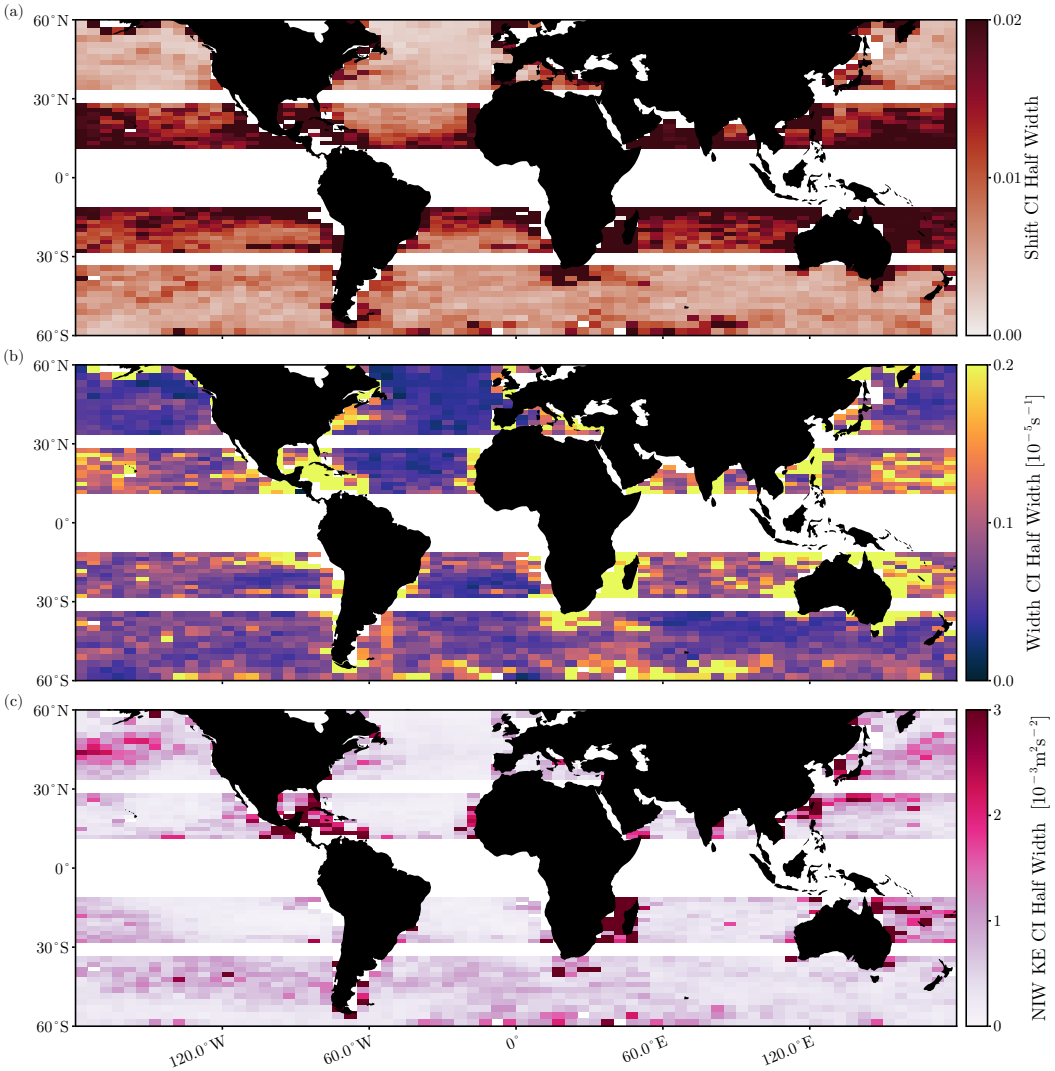

Figure 4.12: 95% confidence interval half width for the characteristics of the NIW peak: (a) the frequency shift ω_I/f, (b) the spectral width σ_I, and (c) the NIW kinetic energy K.

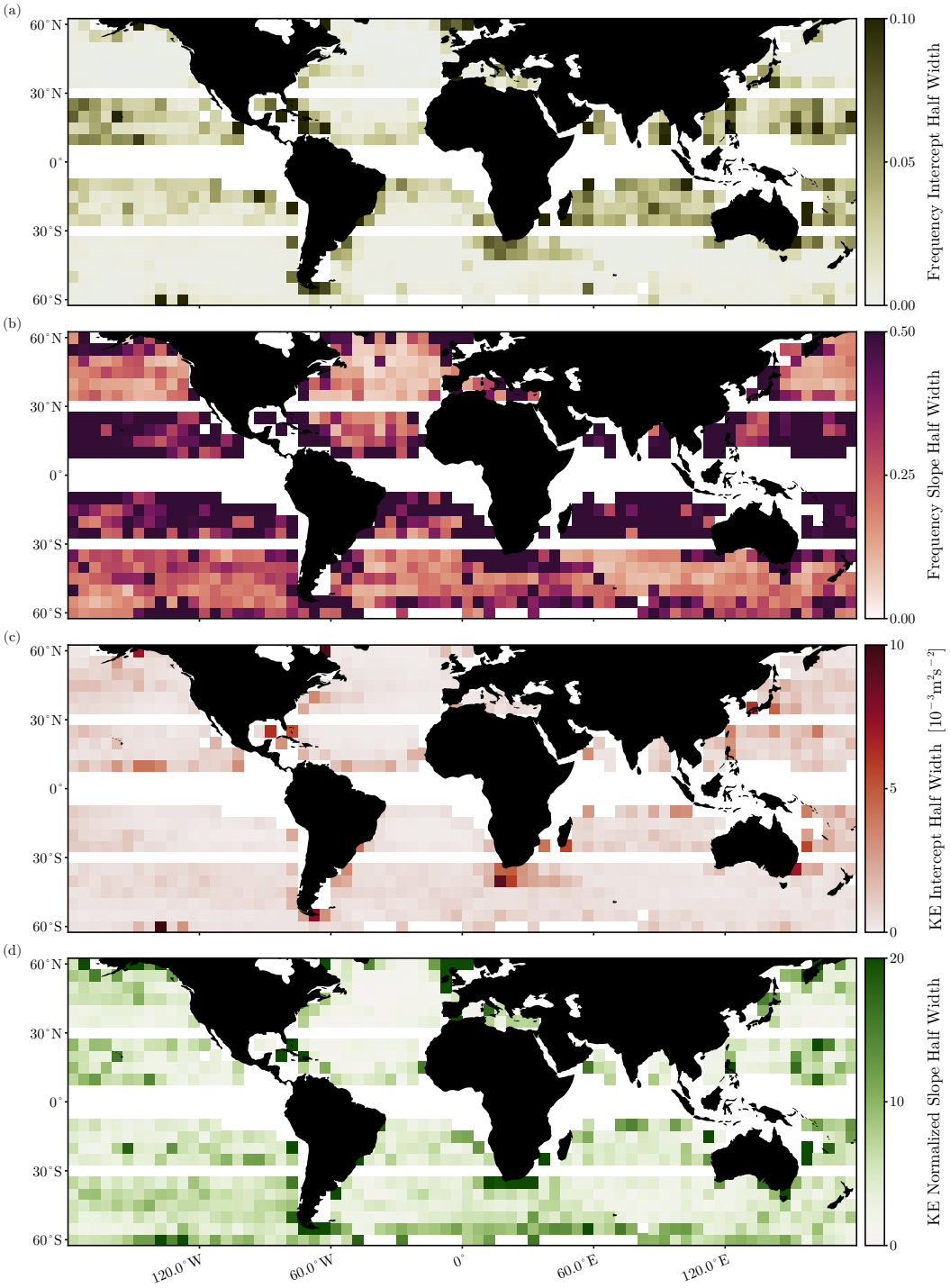

Figure 4.13: 95% confidence interval half width for the geography of the vorticity dependence of NIW spectral properties. (a) Intercept and (b) slope of the linear fit to the NIW frequency shift ω_I / f as a function of vorticity ζ. (c) Intercept and (d) slope of the linear fit to the NIW kinetic energy K as a function of vorticity ζ. Note that we show the slope d normalized by the intercept c because otherwise it primarily reflects the patterns seen in panel (c).

Chapter 5

CONCLUSIONS

The beauty of the YBJ equation lies not in its theoretical elegance, but in its ability to help us understand how NIWs behave in the real world. For instance, consider how the YBJ equation allowed us to ascribe the observed evolution of NIWs to specific physical processes. Trying to glean such information from the observations alone, without the theoretical scaffolding of the YBJ equation, would have been much more difficult. Similarly, the observed NIW spectra from drifters are messy, showing a great deal of variability in their properties across the ocean. The YBJ equation reduces this complexity to a few key parameters, and provides concrete, physically interpretable mechanisms behind this variability. In all, this thesis represents an effort to use theoretical tools to understand the great complexity of NIWs in the real ocean.

Perhaps the most frustrating aspect of NIWs is the difficulty in observing them. Satellite oceanography, which has revolutionised our understanding of mesoscale eddies and tides, is unable to provide similar benefits to our understanding of NIWs. Much of this thesis has been dedicated to using theoretical tools to bridge the gaps in our observations of NIWs. While our understanding of NIW dynamics matures, we still have much to learn about the impacts of this dynamics. For example, we have no observational estimates of the NIW-generated mixing in the upper ocean [1]. In essence, the propagation and death of NIWs remain challengingly elusive. Future work must connect the phenomenology of NIW dynamics to their impacts on the global climate system.

Notably lacking in this thesis was a consideration of the roles of baroclinicity in mesoscale eddies, and submesoscale features in the evolution of NIWs. Assuming barotropic eddies was surprisingly fruitful in understanding NIW dynamics in the upper ocean. The YBJ equation was able to capture the OSMOSIS observations even without knowledge of the baroclinicity of the mesoscale eddy field. Similarly, baroclinicity did not seem necessary to understand the drifter observations. However, when considering the propagation of NIWs to depth, and the roles they can play in the deep ocean, baroclinicity gives rise to a wide variety of new phenomena. In the future, theoretical progress should be made by relaxing the assumption of a barotropic mean flow. Similarly, the heavily smoothed mesoscale eddy field from vorticity seemed adequate to capture the main effects of NIW-mesoscale interactions. But one could imagine circumstances where interaction with smaller-scale vorticity features, down to the submesoscale, could become important. For these scales we have no

[1] Of course, this can be estimated from models but sources of uncertainty are numerous. Regardless, without an observational baseline for comparison, trusting such model estimates is difficult.

such unified theoretical understanding as with the mesoscale. Again, this likely opens up the possibility for new dynamics.

References

, ????: Statistics of simulated and observed pair separations in the Gulf of Mexico, author = Beron-Vera, Francisco J. and LaCasce, J. H., year = 2016, month = jul, journal = Journal of Physical Oceanography, volume = 46, number = 7, pages = 2183–2199, publisher = American Meteorological Society, doi = 10.1175/JPO-D-15-0127.1,.

Alford, M. H., 2001: Internal swell generation: The spatial distribution of energy flux from the wind to mixed layer near-inertial motions. *Journal of Physical Oceanography*, **31 (8)**, 2359–2368.

Alford, M. H., 2020: Revisiting near-inertial wind work: Slab models, relative stress, and mixed layer deepening. *Journal of Physical Oceanography*, **50 (11)**, 3141–3156.

Alford, M. H., A. Le Boyer, A. S. Ren, G. Voet, C. Bellerjeau, C. B. Whalen, B. Hall, and N. Couto, 2025: Observations of turbulence generated by a near-inertial wave propagating downward in an anticyclonic eddy. *Geophysical Research Letters*, **52 (6)**, e2024GL114 070, doi:10.1029/2024GL114070.

Alford, M. H., J. A. MacKinnon, H. L. Simmons, and J. D. Nash, 2016: Near-inertial internal gravity waves in the ocean. *Annual Review of Marine Science*, **8**, 95–123.

Alford, M. H., and M. Whitmont, 2007: Seasonal and spatial variability of near-inertial kinetic energy from historical moored velocity records. *Journal of Physical Oceanography*, **37 (8)**, 2022–2037.

Anderson, D. L., and A. E. Gill, 1979: Beta dispersion of inertial waves. *Journal of Geophysical Research: Oceans*, **84 (C4)**, 1836–1842.

Arakawa, A., 1966: Computational design for long term integration of the equations of motion: Two-dimensional incompressible flow. part i. *Journal of Computational Physics*, **1 (1)**, 9–143.

Arbic, B. K., and G. R. Flierl, 2004: Baroclinically unstable geostrophic turbulence in the limits of strong and weak bottom ekman friction: Application to midocean eddies. *Journal of Physical Oceanography*, **34 (10)**, 2257–2273.

Asselin, O., L. N. Thomas, W. R. Young, and L. Rainville, 2020: Refraction and straining of near-inertial waves by barotropic eddies. *Journal of Physical Oceanography*, **50 (12)**, 3439–3454.

Asselin, O., and W. R. Young, 2019: An improved model of near-inertial wave dynamics. *Journal of Fluid Mechanics*, **876**, 428–448, doi:10.1017/jfm.2019.557.

Asselin, O., and W. R. Young, 2020: Penetration of wind-generated near-inertial waves into a turbulent ocean. *Journal of Physical Oceanography*, **50 (6)**, 1699–1716.

Balmforth, N., S. G. Llewellyn Smith, and W. Young, 1998: Enhanced dispersion of near-inertial waves in an idealized geostrophic flow. *Journal of Marine Researchh*, **56 (1)**, 1–40.

Balmforth, N. J., and W. R. Young, 1999: Radiative damping of near-inertial oscillations in the mixed layer. *Journal of Marine Researchh*, **57 (4)**, 561–584.

Bender, C. M., and S. A. Orszag, 1999: *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory*, Vol. 1, chap. 10. Springer Science & Business Media.

Berry, M., and A. Ozorio de Almeida, 1973: Semiclassical approximation of the radial equation with two-dimensional potentials. *Journal of Physics A: Mathematical, Nuclear and General*, **6 (10)**, 1451.

Berry, M. V., and K. Mount, 1972: Semiclassical approximations in wave mechanics. *Reports on Progress in Physics*, **35 (1)**, 315.

Brillouin, L., 1926: Remarques sur la mécanique ondulatoire. *Journal de Physique et Le Radium*, **7 (12)**, 353–368.

Buckingham, C. E., and Coauthors, 2016: Seasonality of submesoscale flows in the ocean surface boundary layer. *Geophysical Research Letters*, **43 (5)**, 2118–2126.

Bühler, O., 2006: *A brief introduction to classical, statistical, and quantum mechanics*. American Mathematical Society.

Burns, K. J., G. M. Vasil, J. S. Oishi, D. Lecoanet, and B. P. Brown, 2020: Dedalus: A flexible framework for numerical simulations with spectral methods. *Physical Review Research*, **2**, 023 068.

Byun, S.-S., J. J. Park, K.-I. Chang, and R. W. Schmitt, 2010: Observation of near-inertial wave reflections within the thermostad layer of an anticyclonic mesoscale eddy. *Geophysical Research Letters*, **37 (1)**.

Callies, J., R. Barkan, and A. N. Garabato, 2020: Time scales of submesoscale flow inferred from a mooring array. *Journal of Physical Oceanography*, **50 (4)**, 1065–1086.

Chelton, D. B., R. A. DeSzoeke, M. G. Schlax, K. El Naggar, and N. Siwertz, 1998: Geographical variability of the first baroclinic rossby radius of deformation. *Journal of Physical Oceanography*, **28 (3)**, 433–460.

Chelton, D. B., M. G. Schlax, and R. M. Samelson, 2011: Global observations of nonlinear mesoscale eddies. *Progress in Oceanography*, **91 (2)**, 167–216, doi:10.1016/j.pocean.2011.01.002.

Chen, G., H. Xue, D. Wang, and Q. Xie, 2013: Observed near-inertial kinetic energy in the northwestern south china sea. *Journal of Geophysical Research: Oceans*, **118 (10)**, 4965–4977.

Conn, S., J. Callies, and A. Lawrence, 2025: Regimes of near-inertial wave dynamics. *Journal of Fluid Mechanics*, **1002**, A22.

Conn, S., J. Fitzgerald, and J. Callies, 2024: Interpreting observed interactions between near-inertial waves and mesoscale eddies. *Journal of Physical Oceanography*, **54 (2)**, 485–502.

Damerell, G. M., K. J. Heywood, D. Calvert, A. L. Grant, M. J. Bell, and S. E. Belcher, 2020: A comparison of five surface mixed layer models with a year of observations in the north atlantic. *Progress in Oceanography*, **187**, 102 316.

Damerell, G. M., K. J. Heywood, A. F. Thompson, U. Binetti, and J. Kaiser, 2016: The vertical structure of upper ocean variability at the porcupine abyssal plain during 2012–2013. *Journal of Geophysical Research: Oceans*, **121 (5)**, 3075–3089.

Danioux, E., J. Vanneste, and O. Bühler, 2015: On the concentration of near-inertial waves in anticyclones. *Journal of Fluid Mechanics*, **773**, R2.

D'Asaro, E. A., 1985: The energy flux from the wind to near-inertial motions in the surface mixed layer. *Journal of Physical Oceanography*, **15 (8)**, 1043–1059.

D'Asaro, E. A., 1995: Upper-ocean inertial currents forced by a strong storm. Part III: Interaction of inertial currents and mesoscale eddies. *Journal of Physical Oceanography*, **25 (11)**, 2953–2958.

D'Asaro, E. A., C. C. Eriksen, M. D. Levine, P. Niiler, P. Van Meurs, and Coauthors, 1995: Upper-ocean inertial currents forced by a strong storm. Part I: Data and comparisons with linear theory. *Journal of Physical Oceanography*, **25 (11)**, 2909–2936.

Delandmeter, P., and E. Van Sebille, 2019: The parcels v2. 0 lagrangian framework: New field interpolation schemes. *Geoscientific Model Development*, **12 (8)**, 3571–3584.

Dewar, W. K., and G. R. Flierl, 1987: Some effects of the wind on rings. *Journal of Physical Oceanography*, **17 (10)**, 1653–1667.

Edelman, A., and N. R. Rao, 2005: Random matrix theory. *Acta Numerica*, **14**, 233–297.

Edelman, A., and B. D. Sutton, 2007: From random matrices to stochastic operators. *Journal of Statistical Physics*, **127**, 1121–1165.

Einstein, A., 1917: Zum Quantensatz von Sommerfeld and Epstein. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, **19**, 82–92.

Elipot, S., R. Lumpkin, and G. Prieto, 2010: Modification of inertial oscillations by the mesoscale eddy field. *Journal of Geophysical Research: Oceans*, **115 (C9)**.

Essink, S., E. Kunze, R.-C. Lien, R. Inoue, and S.-i. Ito, 2022: Near-inertial wave interactions and turbulence production in a kuroshio anticyclonic eddy. *Journal of Physical Oceanography*, **52 (11)**, 2687–2704.

Fer, I., 2014: Near-inertial mixing in the central arctic ocean. *Journal of Physical Oceanography*, **44 (8)**, 2031–2049.

Ferrari, R., and C. Wunsch, 2009: Ocean circulation kinetic energy: Reservoirs, sources, and sinks. *Annual Review of Fluid Mechanics*, **41**, 253–282.

Flexas, M. M., A. F. Thompson, H. S. Torres, P. Klein, J. T. Farrar, H. Zhang, and D. Menemenlis, 2019: Global estimates of the energy transfer from the wind to the ocean, with emphasis on near-inertial oscillations. *Journal of Geophysical Research: Oceans*, **124 (8)**, 5723–5746.

Forget, G., J.-M. Campin, P. Heimbach, C. Hill, R. Ponte, and C. Wunsch, 2015: Ecco version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, **8 (10)**, 3071–3104.

Fukumori, I., O. Wang, I. Fenty, G. Forget, P. Heimbach, and R. Ponte, 2020: Synopsis of the ecco central production global ocean and sea-ice state estimate, version 4 release 4. zenodo.

Garrett, C., 2001: What is the "near-inertial" band and why is it different from the rest of the internal wave spectrum? *Journal of Physical Oceanography*, **31 (4)**, 962–971.

Gill, A., 1984: On the behavior of internal waves in the wakes of storms. *Journal of Physical Oceanography*, **14 (7)**, 1129–1151.

Gnanadesikan, A., M.-A. Pradal, and R. Abernathey, 2015: Isopycnal mixing by mesoscale eddies significantly impacts oceanic anthropogenic carbon uptake. *Geophysical Research Letters*, **42 (11)**, 4249–4255.

Guan, S., W. Zhao, J. Huthnance, J. Tian, and J. Wang, 2014: Observed upper ocean response to typhoon megi (2010) in the northern south china sea. *Journal of Geophysical Research: Oceans*, **119 (5)**, 3134–3157.

Gutiérrez Brizuela, N., M. H. Alford, S.-P. Xie, J. Sprintall, G. Voet, S. J. Warner, K. Hughes, and J. N. Moum, 2023: Prolonged thermocline warming by near-inertial internal waves in the wakes of tropical cyclones. *Proceedings of the National Academy of Sciences*, **120 (26)**, e2301664 120.

Gutzwiller, M. C., 1992: Quantum chaos. *Scientific American*, **266 (1)**, 78–85.

Hersbach, H., and Coauthors, 2018: Era5 hourly data on single levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, **10 (10.24381)**.

Huang, R., 2004: Energy flows in the ocean. encyclopedia of energy. Elsevier Inc.

Jaimes, B., and L. K. Shay, 2010: Near-inertial wave wake of hurricanes Katrina and Rita over mesoscale oceanic eddies. *Journal of Physical Oceanography*, **40 (6)**, 1320–1337.

Jayne, S. R., and J. Marotzke, 2002: The oceanic eddy heat transport. *Journal of Physical Oceanography*, **32 (12)**, 3328–3345.

Jochum, M., B. P. Briegleb, G. Danabasoglu, W. G. Large, N. J. Norton, S. R. Jayne, M. H. Alford, and F. O. Bryan, 2013: The impact of oceanic near-inertial waves on climate. *Journal of Climate*, **26 (9)**, 2833–2844.

Kafiabad, H. A., J. Vanneste, and W. R. Young, 2021: Interaction of near-inertial waves with an anticyclonic vortex. *Journal of Physical Oceanography*, **51 (6)**, 2035–2048.

Kato, H., and O. Phillips, 1969: On the penetration of a turbulent layer into stratified fluid. *Journal of Fluid Mechanics*, **37 (4)**, 643–655.

Kawaguchi, Y., S. Nishino, J. Inoue, K. Maeno, H. Takeda, and K. Oshima, 2016: Enhanced diapycnal mixing due to near-inertial internal waves propagating through an anticyclonic eddy in the ice-free chukchi plateau. *Journal of Physical Oceanography*, **46 (8)**, 2457–2481.

Keller, J. B., 1958: Corrected Bohr–Sommerfeld quantum conditions for nonseparable systems. *Annals of Physics*, **4 (2)**, 180–188.

Keller, J. B., 1985: Semiclassical mechanics. *Siam Review*, **27 (4)**, 485–504.

Klein, P., S. L. Smith, and G. Lapeyre, 2004: Organization of near-inertial energy by an eddy field. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, **130 (598)**, 1153–1166.

Kunze, E., 1985: Near-inertial wave propagation in geostrophic shear. *Journal of Physical Oceanography*, **15 (5)**, 544–565.

Kunze, E., and T. B. Sanford, 1984: Observations of near-inertial waves in a front. *Journal of Physical Oceanography*, **14 (3)**, 566–581.

Kunze, E., R. W. Schmitt, and J. M. Toole, 1995: The energy balance in a warm-core ring's near-inertial critical layer. *Journal of Physical Oceanography*, **25 (5)**, 942–957.

Large, W., and S. Pond, 1981: Open ocean momentum flux measurements in moderate to strong winds. *Journal of Physical Oceanography*, **11 (3)**, 324–336.

Lazaneo, C. Z., L. Thomas, Z. B. Szuts, J. M. Cusack, K.-F. Chang, and R. K. Shearman, 2024: Interaction of typhoon-driven near-inertial waves with an anticyclone in the philippine sea. *Oceanography*, **37 (4)**, 68–81.

Lee, D.-K., and P. P. Niiler, 1998: The inertial chimney: The near-inertial energy drainage from the ocean surface to the deep layer. *Journal of Geophysical Research: Oceans*, **103 (C4)**, 7579–7591.

Llewellyn Smith, S. G., 1999: Near-inertial oscillations of a barotropic vortex: Trapped modes and time evolution. *Journal of Physical Oceanography*, **29 (4)**, 747–761.

Lucas, A. J., G. C. Pitcher, T. A. Probyn, and R. M. Kudela, 2014: The influence of diurnal winds on phytoplankton dynamics in a coastal upwelling system off southwestern Africa. *Deep Sea Research Part II: Topical Studies in Oceanography*, **101**, 50–62.

Martínez-Marrero, A., and Coauthors, 2019: Near-inertial wave trapping near the base of an anticyclonic mesoscale eddy under normal atmospheric conditions. *Journal of Geophysical Research: Oceans*, **124 (11)**, 8455–8467.

Maslov, V., 1972: *Théorie des perturbations et méthodes asymptotiques*. Dunod.

Middleton, J. F., and C. Garrett, 1986: A kinematic analysis of polarized eddy fields using drifter data. *Journal of Geophysical Research: Oceans*, **91 (C4)**, 5094–5102.

Moehlis, J., and S. G. L. Smith, 2001: Radiation of mixed layer near-inertial oscillations into the ocean interior. *Journal of Physical Oceanography*, **31 (6)**, 1550–1560.

Munk, W., and N. Phillips, 1968: Coherence and band structure of inertial motion in the sea. *Reviews of Geophysics*, **6 (4)**, 447–472, doi:10.1029/RG006i004p00447.

Munk, W., and C. Wunsch, 1998: Abyssal recipes ii: Energetics of tidal and wind mixing. *Deep Sea Research Part I: Oceanographic Research Papers*, **45 (12)**, 1977–2010.

Müller, P., J. McWilliams, and J. Molemaker, 2011: Routes to dissipation in the ocean: The 2D/3D turbulence conundrum. *Marine Turbulence: Theories, Observations, and Models*, Cambridge University Press.

Nikurashin, M., G. K. Vallis, and A. Adcroft, 2013: Routes to energy dissipation for geostrophic flows in the Southern Ocean. *Nature Geoscience*, **6 (1)**, 48–51.

Oey, L.-Y., M. Inoue, R. Lai, X.-H. Lin, S. Welsh, and L. Rouse Jr, 2008: Stalling of near-inertial waves in a cyclone. *Geophysical Research Letters*, **35 (12)**.

Park, J. J., K. Kim, and B. A. King, 2005: Global statistics of inertial motions. *Geophysical research letters*, **32 (14)**.

Park, J. J., K. Kim, and R. W. Schmitt, 2009: Global distribution of the decay timescale of mixed layer inertial motions observed by satellite-tracked drifters. *Journal of Geophysical Research: Oceans*, **114 (C11)**.

Percival, I., and N. Pomphrey, 1976: Vibrational quantization of polyatomic molecules. *Molecular Physics*, **31 (1)**, 97–114.

Percival, I. C., 1977: Semiclassical theory of bound states. *Advances in Chemical Physics*, **36**, 1–61.

Perkins, H., 1976: Observed effect of an eddy on inertial oscillations. *Deep Sea Research and Oceanographic Abstracts*, Elsevier, Vol. 23, 1037–1042.

Pinkel, R., 2008: The wavenumber–frequency spectrum of vortical and internal-wave shear in the western arctic ocean. *Journal of Physical Oceanography*, **38 (2)**, 277–290.

Plueddemann, A., and J. Farrar, 2006: Observations and models of the energy flux from the wind to mixed-layer inertial currents. *Deep Sea Research Part II: Topical Studies in Oceanography*, **53 (1-2)**, 5–30.

Pollard, R. T., and R. Millard Jr, 1970: Comparison between observed and simulated wind-generated inertial oscillations. *Deep Sea Research and Oceanographic Abstracts*, Elsevier, Vol. 17, 813–821.

Pollard, R. T., P. B. Rhines, and R. O. Thompson, 1973: The deepening of the wind-mixed layer. *Geophysical Fluid Dynamics*, **4 (4)**, 381–404.

Qu, L., L. N. Thomas, and R. D. Hetland, 2021: Near-inertial-wave critical layers over sloping bathymetry. *Journal of Physical Oceanography*, **51 (6)**, 1737–1756.

Qu, L., and Coauthors, 2022: Rapid vertical exchange at fronts in the northern gulf of mexico. *Nature Communications*, **13 (1)**, 5624.

Raja, K. J., M. C. Buijsman, J. F. Shriver, B. K. Arbic, and O. Siyanbola, 2022: Near-inertial wave energetics modulated by background flows in a global model simulation. *Journal of Physical Oceanography*, **52 (5)**, 823–840.

Rama, J., C. J. Shakespeare, and A. M. Hogg, 2022a: Importance of background vorticity effect and doppler shift in defining near-inertial internal waves. *Geophysical Research Letters*, **49 (22)**, e2022GL099 498.

Rama, J., C. J. Shakespeare, and A. M. Hogg, 2022b: The wavelength dependence of the propagation of near-inertial internal waves. *Journal of Physical Oceanography*, **52 (10)**, 2493–2514.

Renault, L., M. J. Molemaker, J. C. McWilliams, A. F. Shchepetkin, F. Lemarié, D. Chelton, S. Illig, and A. Hall, 2016: Modulation of wind work by oceanic current interaction with the atmosphere. *Journal of Physical Oceanography*, **46 (6)**, 1685–1704.

Rocha, C. B., G. L. Wagner, and W. R. Young, 2018: Stimulated generation: Extraction of energy from balanced flow by near-inertial waves. *Journal of Fluid Mechanics*, **847**, 417–451.

Rolland, R., and Coauthors, 2025: Near-inertial wave trapping inside a fine-scale anticyclonic eddy during the bioswot-med 2023 cruise: turbulence and energy flux. *ESS Open Archive*.

Sakurai, J. J., and J. Napolitano, 2020: *Modern quantum mechanics*. Cambridge University Press.

Sallée, J.-B., and Coauthors, 2021: Summertime increases in upper-ocean stratification and mixed-layer depth. *Nature*, **591 (7851)**, 592–598.

Sanford, T. B., B. B. Ma, and M. H. Alford, 2021: Stalling and dissipation of a near-inertial wave (niw) in an anticyclonic ocean eddy: Direct determination of group velocity and comparison with theory. *Journal of Geophysical Research: Oceans*, **126 (5)**, e2020JC016 742.

Smith, K. S., 2007: The geography of linear baroclinic instability in Earth's oceans. *Journal of Marine Researchh*, **65 (5)**, 655–683.

Stone, A. D., 2005: Einstein's unknown insight and the problem of quantizing chaos. *Physics Today*, **58 (8)**, 37.

Taburet, G., A. Sanchez-Roman, M. Ballarotta, M.-I. Pujol, J.-F. Legeais, F. Fournier, Y. Faugere, and G. Dibarboure, 2019: Duacs dt2018: 25 years of reprocessed sea level altimetry products. *Ocean Science*, **15 (5)**, 1207–1224.

Thomas, J., K. S. Smith, and O. Bühler, 2017: Near-inertial wave dispersion by geostrophic flows. *Journal of Fluid Mechanics*, **817**, 406–438.

Thomas, L., S. Kelly, T. Klenz, W. Young, L. Rainville, H. Simmons, V. Hormann, and I. Stokes, 2024a: Why near-inertial waves are less affected by vorticity in the northeast pacific than in the north atlantic. *Oceanography*.

Thomas, L. N., 2017: On the modifications of near-inertial waves at fronts: Implications for energy transfer across scales. *Ocean Dynamics*, **67 (10)**, 1335–1350.

Thomas, L. N., J. Moum, L. Qu, J. Hilditch, E. L. Kunze, L. Rainville, and C. Lee, 2024b: Blocked drainpipes and smoking chimneys: Discovery of new near-inertial wave phenomena in anticyclones. *Oceanography*, **37**, URL https://doi.org/10.5670/oceanog.2024.304.

Thomas, L. N., L. Rainville, O. Asselin, W. R. Young, J. Girton, C. B. Whalen, L. Centurioni, and V. Hormann, 2020: Direct observations of near-inertial wave ζ-refraction in a dipole vortex. *Geophysical Research Letters*, **47 (21)**, e2020GL090 375.

Thompson, A. F., A. Lazar, C. Buckingham, A. C. N. Garabato, G. M. Damerell, and K. J. Heywood, 2016: Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders. *Journal of Physical Oceanography*, **46 (4)**, 1285–1307.

Voet, G., A. F. Waterhouse, A. Savage, E. Kunze, and J. A. MacKinnon, 2024: Near-inertial energy variability in a strong mesoscale eddy field in the Iceland Basin. *Oceanography*, **37 (4)**, EOR, doi:10.5670/oceanog.2024.302.

von Storch, J.-S., and V. Lüschow, 2023: Wind power input to ocean near-inertial waves diagnosed from a 5-km global coupled atmosphere-ocean general circulation model. *Journal of Geophysical Research: Oceans*, **128 (2)**, e2022JC019 111.

Wagner, G., and W. Young, 2015: Available potential vorticity and wave-averaged quasi-geostrophic flow. *Journal of Fluid Mechanics*, **785**, 401–424.

Wagner, G., and W. Young, 2016: A three-component model for the coupled evolution of near-inertial waves, quasi-geostrophic flow and the near-inertial second harmonic. *Journal of Fluid Mechanics*, **802**, 806–837.

Whitt, D. B., and L. N. Thomas, 2013: Near-inertial waves in strongly baroclinic currents. *Journal of Physical Oceanography*, **43 (4)**, 706–725.

Wunsch, C., and R. Ferrari, 2004: Vertical mixing, energy, and the general circulation of the oceans. *Annual Review of Fluid Mechanics*, **36**, 281–314.

Xie, J.-H., and J. Vanneste, 2015: A generalised-lagrangian-mean model of the interactions between near-inertial waves and mean flow. *Journal of Fluid Mechanics*, **774**, 143–169.

Young, W., and M. Ben-Jelloul, 1997: Propagation of near-inertial oscillations through a geostrophic flow. *Journal of Marine Researchh*, **55 (4)**, 735–766.

Yu, X., A. C. Naveira Garabato, C. Vic, J. Gula, A. C. Savage, J. Wang, A. F. Waterhouse, and J. A. MacKinnon, 2022a: Observed equatorward propagation and chimney effect of near-inertial waves in the midlatitude ocean. *Geophysical Research Letters*, **49 (13)**, e2022GL098 522.

Yu, X., A. C. Naveira Garabato, C. Vic, J. Gula, A. C. Savage, J. Wang, A. F. Waterhouse, and J. A. MacKinnon, 2022b: Observed equatorward propagation and chimney effect of near-inertial waves in the midlatitude ocean. *Geophysical Research Letters*, **49 (13)**, e2022GL098 522.

Yu, X., A. L. Ponte, S. Elipot, D. Menemenlis, E. D. Zaron, and R. Abernathey, 2019: Surface kinetic energy distributions in the global oceans from a high-resolution numerical model and surface drifter observations. *Geophysical Research Letters*, **46 (16)**, 9757–9766.

Yuan, S., X. Yan, L. Zhang, C. Pang, and D. Hu, 2024: Observation of near-inertial waves induced by typhoon lan in the northwestern pacific: Characteristics, energy fluxes and impact on diapycnal mixing. *Journal of Geophysical Research: Oceans*, **129 (2)**, e2023JC020 187.

Zhai, X., H. L. Johnson, and D. P. Marshall, 2010: Significant sink of ocean-eddy energy near western boundaries. *Nature Geoscience*, **3 (9)**, 608–612.

Zhang, F., and J.-H. Xie, 2023: Scale dependence of near-inertial wave's concentration in anticyclones. *Frontiers in Marine Science*, **10**, 1085 679.