Interactions between Meso- and Sub-Mesoscale Gravity Waves and their Efficient Representation in Mesoscale-Resolving Models

Jannik Wilhelm

Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt, Frankfurt am Main, Germany

T. R. Akylas

Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Gergely Bölöni, Junhong Wei

Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt, Frankfurt am Main, Germany

Bruno Ribstein

CMLA, ENS Cachan, CNRS, Université Paris-Saclay, Cachan, France

CEA, DAM, DIF, Arpajon, France

Rupert Klein

Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

Ulrich Achatz*

Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt, Frankfurt am Main, Germany

Corresponding author address: Institut für Atmosphäre und Umwelt, Goethe Universität Frankfurt, Altenhöferallee 1, 60438 Frankfurt am Main, Germany.

18 E-mail: achatz@iau.uni-frankfurt.de

ABSTRACT

As present weather-forecast codes and increasingly many atmospheric climate models resolve at least part of the mesoscale flow, and hence also internal gravity waves (GWs), it is natural to ask whether even in such configurations sub-gridscale GWs might impact the resolved flow, and how their effect could be taken into account. This motivates a theoretical and numerical investigation of the interaction between unresolved sub-mesoscale and resolved mesoscale GWs, using Boussinesq dynamics for simplicity. By scaling arguments, first a subset of sub-mesoscale GWs that can indeed influence the dynamics of mesoscale GWs is identified. Therein, hydrostatic GWs with wavelengths corresponding to the largest unresolved scales of present-day limited-area weather forecast models are an interesting example. A large-amplitude WKB theory, allowing for a mesoscale unbalanced flow, is then formulated, based on multi-scale asymptotic analysis utilizing a proper scale-separation parameter. Purely vertical propagation of sub-mesoscale GWs is found to be most important, implying inter alia that the resolved flow is only affected by the vertical flux convergence of sub-mesoscale horizontal momentum at leading order. In turn, sub-mesoscale GWs are refracted by mesoscale vertical wind shear while conserving their wave-action density. An efficient numerical implementation of the theory uses a phase-space ray tracer, thus handling the frequent appearance of caustics. The WKB approach and its numerical implementation are validated successfully against sub-mesoscale resolving simulations of the resonant radiation of mesoscale inertia GWs by a horizontally as well as vertically confined sub-mesoscale GW packet.

1. Introduction

Internal gravity waves (GWs) play a significant role in atmospheric dynamics on various spatial scales (Fritts and Alexander 2003; Kim et al. 2003; Alexander et al. 2010; Plougonven and Zhang 2014). Already in the lower atmosphere GW effects are manifold. Examples include the triggering of high-impact weather (e.g. Zhang et al. 2001, 2003) and clear-air turbulence (Koch et al. 2005), as well as the effect of small-scale GWs of orographic origin on the predicted larger-scale flow (e.g. Palmer et al. 1986; Lott and Miller 1997; Scinocca and McFarlane 2000) and the GW impact on the generation of high cirrus clouds and polar stratospheric clouds (e.g. Joos et al. 2009). Even more conspicuous than in the lower atmosphere, however, are GW effects in the middle atmosphere. The general circulation in the mesosphere is basically controlled by GWs (Lindzen 1981; Holton 1982; Garcia and Solomon 1985). This also seems to be of relevance to both medium-range weather forecast and climate modeling in the troposphere. Middle-atmosphere circulation influences the lower layers by downward control (Haynes et al. 1991), and there is evidence of the importance of the middle atmosphere for long-range forecasting of winter weather (Baldwin and Dunkerton 2001; Kidston et al. 2015; Hansen et al. 2017; Jia et al. 2017) and climate (Scaife et al. 2005, 2012) in the northern hemisphere.

As a substantial portion of the GW spectrum involves scales too small to describe explicitly in current-resolution climate models, accounting for such small-scale GWs poses an important parameterization problem to atmospheric dynamics. With rising computing power available, an increasing number of studies of middle-atmosphere global GW dynamics uses models that can resolve a part of the GW spectrum (Kawatani et al. 2009, 2010a,b; Brune and Becker 2013). This raises the question whether the neglected sub-gridscale (SGS) GWs could impact the resolved flow; and, if so, how their effect could be taken into account. In regard to these issues, to the best

of our knowledge, global weather-forecast codes, with horizontal mesh distances of $O(10\,\mathrm{km})$ still generally use a parameterization of SGS GWs, while high-resolution local-area codes used by the weather services, with mesh distances of $O(1\,\mathrm{km})$, typically do not. Since the 'effective resolution' in such codes is well above their mesh distances (Skamarock 2004; Ricard et al. 2013), one might suppose that even there a considerable portion of the GW spectrum is not captured. However, a systematic investigation of the potential impact of SGS GWs on the resolved mesoscale flow is lacking at present.

Available GW parameterizations (e.g. Lindzen 1981; Palmer et al. 1986; McFarlane 1987; Alexander and Dunkerton 1999; Warner and McIntyre 2001; Scinocca 2003; Orr et al. 2010) invariably rely on WKB theory (Bretherton 1966) for describing the interaction between scale-separated waves and (resolved) mean flow. However, the specific implications of this theory may depend on the scales involved. The classic scenario is the interaction between a resolved synoptic-scale flow and unresolved mesoscale inertia GWs. The corresponding WKB theory (Grimshaw 1975; Achatz et al. 2017) as well as the generalized Lagrangian-mean theory (Andrews and McIntyre 1978a,b; Bühler 2009) show that the wave amplitude is controlled by wave-action conservation, while the synoptic-scale flow is described by a quasi-geostrophic potential vorticity that is affected by the GWs via pseudomomentum-flux convergence. For efficiency reasons, parameterizations use these theoretical results with drastic simplifications: (i) lateral GW propagation and the impact of horizontal mean-flow gradients are ignored; and (ii) the time-dependent transient wave mean-flow interaction is replaced by an equilibrium picture where, due to the non-acceleration paradigm, GWs can only modify the resolved flow when they break. In the present context especially the latter steady-state approximation may not be entirely justified. As pointed out by Bühler and McIntyre (1998, 2003, 2005) wave transience is potentially important, and recently Bölöni et al. (2016) have shown that in many cases it can attain at least an equally important role as tur-

bulent wave breaking in mediating the impact of GWs on the resolved flow. It is also essential to keep in mind that the standard WKB approach assumes from the outset geostrophic and hydrostatic balance of the synoptic-scale flow. It is therefore not obvious that this theory can be applied to the interaction between a mesoscale resolved flow and mesoscale or sub-mesoscale SGS GWs, which seems to be the most appropriate scenario for GW parameterizations in mesoscale-resolving models. In this setting, a modified WKB theory that allows for a mesoscale unbalanced large-scale flow, would be most useful.

Of related interest is that packets of small-scale GWs are capable of radiating larger-scale GWs. This possibility was first suggested by Bretherton (1969) for two-dimensional small-scale GW packets with isotropic scaling, and more recently has been investigated further by Van den Bremer and Sutherland (2014) for wave packets of various aspect ratios. The radiation of large-scale waves hinges on a resonance mechanism, wherein the vertical phase velocity of the emitted long waves matches the vertical group velocity of the small-scale wave packet, which acts as a traveling wave source. Furthermore, the vertical wave number of the long wave is set by the scale of the wave packet envelope. In a related study Tabaei and Akylas (2007) show that the long-wave radiation process is especially enhanced if the small-scale wave packet is 'flat' (i.e. its envelope is elongated in the horizontal relative to the vertical) so that both the horizontal and vertical envelope scales can be compatible with free, nearly steady, long GWs. All of these studies assume the small-scale GWs to be non-hydrostatic. They do not investigate, however, which small-scale GWs in general are able to interact with given mesoscale long GWs.

Moreover, no prior study examines the feasibility of a model for SGS GWs in a resolved mesoscale flow. Closest to this comes the WKB model of Tabaei and Akylas (2007). These authors, however, report numerical instability problems once the wave mean-flow interaction develops caustics where the initially locally monochromatic small-scale GW field exhibits multival-

ued wave numbers. This problem, also observed by Rieper et al. (2013a), can be circumvented, however. As shown by Muraschko et al. (2015) and Bölöni et al. (2016), a spectral approach based on phase-space wave-action density yields numerically stable and fast algorithms for the efficient integration of the coupled equations of small-scale GWs in a larger-scale flow.

Building on the above brief review of related prior literature, the goals of the present paper are: (i) a systematic investigation of which smaller-scale GWs are able to interact resonantly with given typical mesoscale GWs; (ii) the development of a WKB theory for the efficient description of this interaction; (iii) the implementation of a numerical algorithm for this theory; and finally (iv) the validation of the WKB theory and its numerical implementation against sub-mesoscale resolving simulations of the radiation of mesoscale GWs by horizontally and vertically confined sub-mesoscale GW packets as considered earlier by Bretherton (1969), Van den Bremer and Sutherland (2014) and Tabaei and Akylas (2007).

The paper is structured as follows. The GW scales of interest are identified in sections 2a and 2b; these scales form the basis for nondimensionalizing the Boussinesq equations and formulating an appropriate multi-scale asymptotic ansatz in section 2c. Leading- and next-order WKB approximations are discussed in section 2d, which eventually yield (section 2e) a coupled energy-conserving equation system of linear Boussinesq equations for the mesoscale, and one-dimensional ray equations for the sub-mesoscale dynamics. Subsequently, our numerical models are described in section 3; the initial conditions of the numerical experiments are motivated in section 4a, and section 4b briefly discusses the postprocessing of the model output data. In section 4c, a kinematic analysis similar to ship wake theory is used to predict the geometry of the induced mesoscale wave disturbance, on the assumption of steady-state forcing by a propagating sub-mesoscale GW packet. In section 4d, the simulation results of different test cases are presented

and compared against the theoretical predictions. Finally, the article concludes with a summary and discussion of the main findings in section 5.

2. Theory - Basics and formalism

For simplicity the interaction between mesoscale and sub-mesoscale GWs is studied in a rotating, incompressible and inviscid Boussinesq atmosphere with height-dependent background stratification, characterized by Coriolis parameter f and Brunt-Väisälä frequency $N(z)$. Under these flow conditions, the governing equations are

$$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} + f\mathbf{e}_z \times \mathbf{u} = -\nabla p + b\mathbf{e}_z \tag{1}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} + N^2 w = 0 \tag{2}$$

$$\nabla \cdot \mathbf{v} = 0\,, \tag{3}$$

where $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \mathbf{v}\cdot\nabla$ is the material derivative, $\mathbf{e}_z$ denotes the unit vector pointing upwards, $\mathbf{v}$ is the full and $\mathbf{u}$ the horizontal velocity vector, while w stands for the vertical velocity component. Furthermore, p and b are the density-weighted pressure deviation and the buoyancy deviation, respectively, from a reference atmosphere with stratification $N^2(z)$, generally slowly varying in the vertical. These equations can capture essential aspects of local dynamics at various scales, as long as the vertical length scale of the waves is smaller than the atmospheric density scale height.

a. Wave scaling

In the following a (resolved) mesoscale flow interacting with (unresolved) smaller-scale motions is considered, termed 'sub-mesoscale' for simplicity. The question arises which, if any, sub-

mesoscale motions are able to leave an impact on the mesoscale flow. Here this issue is addressed by considering possible interactions between a mesoscale and a sub-mesoscale GW.

The mesoscale GW (subscript m) is taken to have horizontal and vertical length scales L_m and H_m, respectively, with an aspect ratio of the order

$$a_m = \frac{H_m}{L_m} = \frac{f}{N_*} \,, \tag{4}$$

where N_* is a characteristic value of the Brunt-Väisälä frequency N, so that the mesoscale intrinsic frequency $\hat{\omega}_m$ is equally affected by rotation and stratification, as follows from the general GW dispersion relation

$$\hat{\omega}^2 = \frac{f^2 m^2 + N^2(k^2+l^2)}{k^2+l^2+m^2} \,. \tag{5}$$

Here, $N^2 = O(N_*^2)$ is meant to be the local value of the stratification, k and l are the horizontal and m the vertical wave vector components, so that $L_m = O(1/\sqrt{k_m^2+l_m^2})$ and $H_m = O(1/m_m)$. These scaling assumptions are met, for instance, in the case where the horizontal and vertical scale are smaller by a synoptic-scale Rossby number than the Rossby deformation radius and the vertical scale height, respectively; in such situations one can take $(H_m, L_m) = (1, 100)$ km (Achatz et al. 2017). The sub-mesoscale wave (subscript w) has shorter vertical and horizontal scales H_w and L_w, respectively, satisfying

$$(H_w, L_w) = (\eta H_m, \eta^p L_m) \tag{6}$$

with $\eta \ll 1$ a small parameter and $p > 0$, so that the sub-mesoscale aspect ratio is of the order

$$a_w = \frac{H_w}{L_w} = \eta^{1-p} \frac{f}{N_*} \,. \tag{7}$$

The respective wave amplitudes are chosen as large as possible while keeping the analysis tractable. Hence the sub-mesoscale wave is assumed to be at the margin of static instability, where $\mathrm{d}b/\mathrm{d}z = O(N^2)$, so that the buoyancy-amplitude scale is

$$B_w = H_w N_*^2 \ , \tag{8}$$

implying vertical displacements of the order of the sub-mesoscale vertical scale. As for the mesoscale flow, if a marginally statically stable buoyancy amplitude is assumed, it turns out that the sub-mesoscale-wave frequency $\omega = \hat{\omega} + \mathbf{k}_{w\mathrm{h}} \cdot \mathbf{u}_m$ - with $\mathbf{k}_{w\mathrm{h}}$ indicating the horizontal sub-mesoscale wave vector and $\mathbf{u}_m$ representing the mesoscale horizontal wind - is dominated by the Doppler term due to the strong mesoscale-flow horizontal winds, while the intrinsic frequency $\hat{\omega}$ is relatively small; as a result, sub-mesoscale motions are mainly transported by the mesoscale flow in such a regime. For this reason, the mesoscale buoyancy-amplitude scale is restricted to satisfy

$$B_m = \eta H_m N_*^2 = B_w \ , \tag{9}$$

that actually agrees with the sub-mesoscale buoyancy-amplitude scale.

The remaining scales of interest follow from the buoyancy equation (2) and the continuity equation (3), with the material derivative scaling with the intrinsic frequency, which provides the inverse time scale. Specifically, the buoyancy equation yields a mesoscale vertical-wind scale:

$$W_m = \frac{f B_m}{N_*^2} = \eta f H_m \ . \tag{10}$$

From the continuity equation follows a mesoscale horizontal-wind scale

$$U_m = \frac{W_m L_m}{H_m} = \eta f L_m \ , \tag{11}$$

and one also notes for later reference that the mesoscale time scale is $T_m = 1/f$. Likewise one obtains for the sub-mesoscale horizontal- and vertical-velocity scales

$$U_w = \frac{W_w L_w}{H_w} = \frac{L_w}{H_w}\frac{\Omega_w}{f} W_m = \frac{L_w}{H_w}\frac{\Omega_w}{N_*} U_m = \eta^{p-1}\frac{\Omega_w}{f} U_m \tag{12}$$

$$W_w = \frac{\Omega_w B_w}{N_*^2} = \Omega_w H_w = \frac{\Omega_w}{f} W_m \ , \tag{13}$$

where Ω_w is the scale-dependent sub-mesoscale intrinsic-frequency scale. From the dispersion relation (5), to a good approximation,

$$\Omega_w(a_w) = \begin{cases} f, & a_w \le \dfrac{f}{N_*} \\ N_* a_w, & \dfrac{f}{N_*} < a_w \le 1 \\ N_*, & a_w > 1 \end{cases} \ . \tag{14}$$

Of these, the last, strongly non-hydrostatic regime is generally modulationally unstable (Sutherland 2001) and hence not considered here.

b. Regimes of interaction between mesoscale and sub-mesoscale motions

Decomposing the total flow into mesoscale and smaller-scale sub-mesoscale motions, $(\mathbf{v}, b) = (\mathbf{v}_m, b_m) + (\mathbf{v}_w, b_w)$, it is expected that the latter can only influence the former via flux-convergence terms. For such an interaction to be possible, these terms must be of the same magnitude as (or larger than) the leading mesoscale terms in the governing equations. To meet this condition,

sub-mesoscale wave fields are considered, with scaling as introduced above, that are spatially modulated on the mesoscale, in response to the two-way interaction between mesoscale and sub-mesoscale flow. As explained in the Appendix, it is then possible to identify a sufficiently scale-separated regime, where sub-mesoscale motions may interact significantly with mesoscale GWs. In this regime the small-scale GWs lie in the mid-frequency range $f/N_* < a_w \leq 1$, and their scale separation is obtained by setting $p = 2$.

Specifically, incorporating this finding in (7) leads to a sub-mesoscale aspect ratio

$$a_w = \frac{f}{\eta N_*} \ . \tag{15}$$

For the mid-frequency range $f/N_* < a_w \leq 1$ this implies, under the requirement of a sufficiently strong scale separation $\eta \ll 1$,

$$\frac{f}{N_*} \leq \eta \ll 1 \ . \tag{16}$$

According to (6), this scaling bears a stronger scale separation in the horizontal than in the vertical

$$(H_w, L_w) = \left(\eta H_m, \eta^2 L_m\right) \ . \tag{17}$$

In keeping with equations (15) and (16), $a_w = \eta^q$ is put, where $q \geq 0$, so that

$$\eta = \left(\frac{f}{N_*}\right)^{\frac{1}{1+q}} \ . \tag{18}$$

While, asymptotically, $q \geq 0$ is a free parameter, for atmospheric applications where $f/N_* = O(10^{-2})$, a sufficiently scale-separated general scaling regime can be identified in the finite range

$0 \leq q \leq 1$. In view of (17) and (18), two characteristic limit cases thus arise

$$q = 0: \quad \eta = \frac{f}{N_*}, \qquad H_w = L_w = \frac{f}{N_*} H_m \tag{19}$$

$$q = 1: \quad \eta = \sqrt{\frac{f}{N_*}}, \quad H_w = \sqrt{\frac{f}{N_*}} H_m, \quad L_w = \frac{f}{N_*} L_m \, . \tag{20}$$

The first (non-hydrostatic) limit case is the one also discussed by Tabaei and Akylas (2007). Here the scale separation is quite large and the mesoscale-wave amplitude that can be affected is rather small. The most interesting case for atmospheric applications is the second (hydrostatic) limit, for which the mesoscale-wave impact is the strongest. For instance, taking $(H_m, L_m) = (1, 100)$ km and $f/N_* = 10^{-2}$, from (20) is then found that $(H_w, L_w) = (0.1, 1)$ km. Notably, this scale estimate is in good agreement with present-day local-area weather-forecast-code mesh distances (see section 1).

For later reference, table 1 provides an overview of the scales deduced in this section. It is worth noting that the pressure scales follow from $\partial p_{m,w}/\partial z = O(b_{m,w})$, which hold both in the hydrostatic and non-hydrostatic regime, so that

$$(P_m, P_w) = (H_m B_m, H_w B_w) = (\eta H_m^2 N_*^2, \eta^2 H_m^2 N_*^2) \, , \tag{21}$$

as can be also verified from the GW polarization relations. Finally, the sub-mesoscale time scale is the inverse intrinsic-frequency scale

$$T_w = \frac{1}{\Omega_w} = \frac{f}{N_* a_w} \frac{1}{f} = \eta T_m \, . \tag{22}$$

c. Nondimensional equations, multi-scale asymptotics and WKB ansatz

In the next step, the scaling for the sub-mesoscale GWs derived above (see table 1) is used to nondimensionalize the governing equations (1) - (3). After substituting

$$(\mathbf{u}, w, b, p) \rightarrow (U_w\mathbf{u}, W_w w, B_w b, P_w p) \tag{23}$$

$$(\mathbf{x}_\mathrm{h}, z, t) \rightarrow (L_w\mathbf{x}_\mathrm{h}, H_w z, T_w t) \tag{24}$$

$$\left(f, N^2\right) \rightarrow \left(f f_0, N_*^2 N_0^2\right) = \left(f f_0, \frac{N_0^2}{a_w^2 T_w^2}\right) , \tag{25}$$

where the subscript h denotes the horizontal components (here of the position vector $\mathbf{x}$), the dimensionless equation system reads

$$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} + \eta f_0 \mathbf{e}_z \times \mathbf{u} = -\nabla_\mathrm{h} p - \frac{1}{\eta^{2q}}\left(\frac{\partial p}{\partial z} - b\right)\mathbf{e}_z \tag{26}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} + N_0^2 w = 0 \tag{27}$$

$$\nabla \cdot \mathbf{v} = 0 . \tag{28}$$

Next, 'compressed' variables are introduced to describe the slow variations of the resolved mesoscale flow, as compared to those of the sub-mesoscale flow,

$$(\mathbf{X}, T) = (\mathbf{X}_\mathrm{h}, Z, T) = \left(\eta^2 \mathbf{x}_\mathrm{h}, \eta z, \eta t\right) , \tag{29}$$

and then a WKB ansatz to describe a locally monochromatic sub-mesoscale wave with slowly varying amplitude, wave number and frequency is used. For a generic variable ξ it reads

$$\xi(\mathbf{x}, t) = \Re\, \Xi(\mathbf{X}, T) \mathrm{e}^{\mathrm{i}\frac{\phi(\mathbf{X},T)}{\eta^2}} , \tag{30}$$

where $\Xi(\mathbf{X},T)$ indicates the (slowly varying) amplitude and $\phi(\mathbf{X},T)\eta^{-2}$ the (rapidly varying) phase. Following Tabaei and Akylas (2007), the latter is defined as $\phi(\mathbf{X},T) = \phi_0(\mathbf{X}_\mathrm{h}) + \eta\,\phi_1(\mathbf{X}_\mathrm{h},Z,T)$, so that the local horizontal wave number, vertical wave number and frequency are

$$\nabla_\mathrm{h}\left(\frac{\phi}{\eta^2}\right) = \nabla_{X_\mathrm{h}}\phi_0 + \eta\,\nabla_{X_\mathrm{h}}\phi_1 \equiv \mathbf{k}_\mathrm{h}^{(0)}(\mathbf{X}_\mathrm{h}) + \eta\,\mathbf{k}_\mathrm{h}^{(1)}(\mathbf{X}_\mathrm{h},Z,T) \tag{31}$$

$$\frac{\partial}{\partial z}\left(\frac{\phi}{\eta^2}\right) = \frac{\partial\phi_1}{\partial Z} \equiv m(\mathbf{X}_\mathrm{h},Z,T) \tag{32}$$

$$\frac{\partial}{\partial t}\left(\frac{\phi}{\eta^2}\right) = \frac{\partial\phi_1}{\partial T} \equiv -\omega(\mathbf{X}_\mathrm{h},Z,T)\ , \tag{33}$$

where ∇_X denotes compressed spatial derivatives. Finally, all fields are expanded in the small-scale separation parameter η, taking into account the scaling derived above, and using subscripts 0 and 1 for the mesoscale and sub-mesoscale parts, respectively:

$$\begin{pmatrix} \mathbf{u} \\ w \\ b \\ p \end{pmatrix} = \sum_{j=0}^{\infty}\eta^j \begin{bmatrix} \mathbf{u}_0^{(j)}(\mathbf{X},T) \\ \eta w_0^{(j)}(\mathbf{X},T) \\ b_0^{(j)}(\mathbf{X},T) \\ \eta^{-1}p_0^{(j)}(\mathbf{X},T) \end{bmatrix} + \Re\sum_{j=0}^{\infty}\eta^j \begin{bmatrix} \mathbf{u}_1^{(j)}(\mathbf{X},T) \\ w_1^{(j)}(\mathbf{X},T) \\ b_1^{(j)}(\mathbf{X},T) \\ p_1^{(j)}(\mathbf{X},T) \end{bmatrix} \mathrm{e}^{\mathrm{i}\frac{\phi(\mathbf{X},T)}{\eta^2}}\ . \tag{34}$$

Note that in this ansatz both mesoscale-field and sub-mesoscale-wave amplitude as well as spatial and temporal scaling are all given in terms of the scale separation parameter η. There is no separate amplitude parameter. Higher harmonics of the sub-mesoscale waves are neglected, as they can be shown to not contribute at leading order due to the dispersive GW dispersion relation (Achatz et al. 2017).

d. Order analysis

After inserting the multi-scale asymptotic ansatz (34) into the nondimensional equations (26) - (28) all terms are sorted by equal powers of η and the phase factor $\mathrm{e}^{\mathrm{i}\eta^{-2}\phi(\mathbf{X},T)}$. Terms without phase factor describe mesoscale dynamics, while those proportional to the phase factor yield information on sub-mesoscale wave dynamics; other harmonics, like $\mathrm{e}^{2\mathrm{i}\eta^{-2}\phi(\mathbf{X},T)}$ and higher, are not considered, as noted above. The following will be kept concise as the procedure is standard (e.g. Achatz et al. 2010, 2017).

1) LEADING ORDER RESULTS

The leading order of the vertical momentum equation establishes that the mesoscale flow is hydrostatic,

$$\frac{\partial p_0^{(0)}}{\partial Z} - b_0^{(0)} = 0 \,, \tag{35}$$

while the leading-order sub-mesoscale terms in the equations can be summarized as $\mathbf{M}_q\left(\mathbf{k}_\mathrm{h}^{(0)}, m, \hat{\omega}\right)\ \mathbf{s}_1^{(0)} = \mathbf{0}$, where $\mathbf{M}_q$ is the anti-hermitian coefficient matrix

$$\mathbf{M}_q\left(\mathbf{k}_\mathrm{h}^{(0)}, m, \hat{\omega}\right) = \begin{pmatrix} -\mathrm{i}\hat{\omega} & 0 & 0 & 0 & \mathrm{i}k^{(0)} \\ 0 & -\mathrm{i}\hat{\omega} & 0 & 0 & \mathrm{i}l^{(0)} \\ 0 & 0 & -\mathrm{i}\hat{\omega}\delta_{q0} & -N_0 & \mathrm{i}m \\ 0 & 0 & N_0 & -\mathrm{i}\hat{\omega} & 0 \\ \mathrm{i}k^{(0)} & \mathrm{i}l^{(0)} & \mathrm{i}m & 0 & 0 \end{pmatrix} \tag{36}$$

with the well-known Kronecker delta δ_{q0}, the intrinsic frequency $\hat{\omega} = \omega - \mathbf{k}_\mathrm{h}^{(0)} \cdot \mathbf{u}_0^{(0)}$, i.e. the frequency relative to the mesoscale velocity, and

$$\mathbf{s}_1^{(0)} = \left(u_1^{(0)}, v_1^{(0)}, w_1^{(0)}, \frac{b_1^{(0)}}{N_0}, p_1^{(0)} \right)^T , \tag{37}$$

the vector of leading-order sub-mesoscale wave amplitudes. Non-trivial sub-mesoscale wave amplitudes $\mathbf{s}_1^{(0)} \neq \mathbf{0}$ require a vanishing determinant of M_q, leading to either the balanced solution $\hat{\omega} = 0$ or the GW dispersion relation

$$\hat{\omega} = \pm N_0 \sqrt{\frac{\left|\mathbf{k}_\mathrm{h}^{(0)}\right|^2}{\left|\mathbf{k}_\mathrm{h}^{(0)}\right|^2 \delta_{q0} + m^2}} . \tag{38}$$

The corresponding null vector yields the sub-mesoscale wave-amplitude polarization relations

$$\left\{ w_1^{(0)}, p_1^{(0)}, \mathbf{u}_1^{(0)} \right\} = \mathrm{i}\hat{\omega} \left\{ \frac{1}{m}, -\frac{\hat{\omega}}{\left|\mathbf{k}_\mathrm{h}^{(0)}\right|^2}, -\frac{\mathbf{k}_\mathrm{h}^{(0)}}{\left|\mathbf{k}_\mathrm{h}^{(0)}\right|^2} \right\} a^{(0)} , \tag{39}$$

where $a^{(0)} = m N_0^{-2} b_1^{(0)}$ defines the buoyancy amplitude relative to the margin of static instability.

While the leading-order horizontal wave number does not develop in time, vertical wave number and frequency do. From their definition and the dispersion relation

$$\omega(\mathbf{X}, T) = \mathbf{k}_\mathrm{h}^{(0)} \cdot \mathbf{u}_0^{(0)}(\mathbf{X}, T) \pm N_0(Z) \sqrt{\frac{\left|\mathbf{k}_\mathrm{h}^{(0)}\right|^2}{\left|\mathbf{k}_\mathrm{h}^{(0)}\right|^2 \delta_{q0} + m^2}} \equiv \Omega(\mathbf{X}, T, \mathbf{k}) , \tag{40}$$

one obtains their prognostic eikonal equations

$$\left(\frac{\partial}{\partial T}+c_{\mathrm{g}z}\frac{\partial}{\partial Z}\right)\omega = \frac{\partial\Omega}{\partial T}=\mathbf{k}_{\mathrm{h}}^{(0)}\cdot\frac{\partial\mathbf{u}_0^{(0)}}{\partial T} \tag{41}$$

$$\left(\frac{\partial}{\partial T}+c_{\mathrm{g}z}\frac{\partial}{\partial Z}\right)m = -\frac{\partial\Omega}{\partial Z}=-\mathbf{k}_{\mathrm{h}}^{(0)}\cdot\frac{\partial\mathbf{u}_0^{(0)}}{\partial Z}\mp\frac{\mathrm{d}N_0}{\mathrm{d}Z}\sqrt{\frac{\left|\mathbf{k}_{\mathrm{h}}^{(0)}\right|^2}{\left|\mathbf{k}_{\mathrm{h}}^{(0)}\right|^2\delta_{q0}+m^2}}\,, \tag{42}$$

where $c_{\mathrm{g}z}=\partial\Omega/\partial m$ is the (intrinsic) vertical group velocity. No horizontal group velocities appear since, at the sub-mesoscales considered, to leading order, energy is transported only vertically.

2) Higher-order results

The next-to-leading ('second') orders in η yield the following mesoscale-flow equations:

$$\frac{\partial b_0^{(0)}}{\partial T}+N_0^2 w_0^{(0)} = 0 \tag{43}$$

$$\frac{\partial\mathbf{u}_0^{(0)}}{\partial T}+f_0\mathbf{e}_z\times\mathbf{u}_0^{(0)}+\nabla_{X_{\mathrm{h}}}p_0^{(0)} = -\frac{1}{2}\frac{\partial}{\partial Z}\Re\left\{w_1^{(0)}\mathbf{u}_1^{(0)*}\right\}\,, \tag{44}$$

while the 'third' order of the mesoscale part of the continuity equation reads

$$\nabla_X\cdot\mathbf{v}_0^{(0)}=0\,, \tag{45}$$

with $\mathbf{v}_0^{(0)}=\mathbf{u}_0^{(0)}+w_0^{(0)}\mathbf{e}_z$. As expected, a sub-mesoscale-wave impact only exists in the horizontal momentum equation.

With the definition

$$\mathbf{r}_q = -\begin{pmatrix} \left(\frac{\partial}{\partial T}+\mathrm{i}\upsilon\right)u_1^{(0)} + w_1^{(0)}\frac{\partial u_0^{(0)}}{\partial Z} - f_0 v_1^{(0)} + \mathrm{i}k^{(1)}p_1^{(0)} \\ \left(\frac{\partial}{\partial T}+\mathrm{i}\upsilon\right)v_1^{(0)} + w_1^{(0)}\frac{\partial v_0^{(0)}}{\partial Z} + f_0 u_1^{(0)} + \mathrm{i}l^{(1)}p_1^{(0)} \\ \left(\frac{\partial}{\partial T}+\mathrm{i}\upsilon\right)w_1^{(0)}\delta_{q0} + \frac{\partial p_1^{(0)}}{\partial Z} \\ \frac{1}{N_0}\left(\frac{\partial}{\partial T}+\mathrm{i}\upsilon\right)b_1^{(0)} + \frac{1}{N_0}w_1^{(0)}\frac{\partial b_0^{(0)}}{\partial Z} \\ \mathrm{i}\mathbf{k}_\mathrm{h}^{(1)}\cdot\mathbf{u}_1^{(0)} + \frac{\partial w_1^{(0)}}{\partial Z} \end{pmatrix}, \tag{46}$$

using the shortcut $\upsilon = \mathbf{k}_\mathrm{h}^{(0)}\cdot\mathbf{u}_0^{(1)} + \mathbf{k}_\mathrm{h}^{(1)}\cdot\mathbf{u}_0^{(0)} + mw_0^{(0)}$, the sub-mesoscale wave terms of the equations give for the next-to-leading orders in η the equation set $\mathbf{M}_q\left(\mathbf{k}_\mathrm{h}^{(0)}, m, \hat{\omega}\right)\ \mathbf{s}_1^{(1)} = \mathbf{r}_q,$ where

$$\mathbf{s}_1^{(1)} = \left(u_1^{(1)}, v_1^{(1)}, w_1^{(1)}, \frac{b_1^{(1)}}{N_0}, p_1^{(1)}\right)^T \tag{47}$$

contains the next-order wave amplitudes. $\mathbf{M}_q$ has a non-vanishing nullspace and thus $\mathbf{r}_q$ may not project onto it. This amounts to $\mathbf{s}_1^{(0)\dagger}\mathbf{r}_q = 0,$ with † for the complex conjugate transpose, yielding

$$\frac{\partial e_w}{\partial T} + \frac{1}{2}\frac{\partial}{\partial Z}\Re\left\{w_1^{(0)}p_1^{(0)*}\right\} = -\frac{1}{2}\Re\left\{w_1^{(0)}\mathbf{u}_1^{(0)*}\right\}\cdot\frac{\partial \mathbf{u}_0^{(0)}}{\partial Z}\ , \tag{48}$$

where * denotes the complex conjugate, for the energy density

$$e_w = e_1^{(0)} = \frac{1}{4}\left(\left|\mathbf{u}_1^{(0)}\right|^2 + \left|w_1^{(0)}\right|^2\delta_{q0} + \frac{1}{N_0^2}\left|b_1^{(0)}\right|^2\right) = \frac{1}{2N_0^2}\left|b_1^{(0)}\right|^2\ . \tag{49}$$

Using the dispersion relation (38) and polarization relations (39), the energy flux and the shear-production term are expressed as

$$\frac{1}{2}\Re\left\{w_1^{(0)}p_1^{(0)*}\right\} = c_{\mathrm{g}z}e_w \tag{50}$$

$$\frac{1}{2}\Re\left\{w_1^{(0)}\mathbf{u}_1^{(0)*}\right\}\cdot\frac{\partial\mathbf{u}_0^{(0)}}{\partial Z} = c_{\mathrm{g}z}\frac{e_w}{\hat{\omega}}\mathbf{k}_{\mathrm{h}}^{(0)}\cdot\frac{\partial\mathbf{u}_0^{(0)}}{\partial Z}\ . \tag{51}$$

Thus, the wave-action conservation equation is obtained from (48), yielding

$$\frac{\partial\mathscr{A}_w}{\partial T}+\frac{\partial}{\partial Z}\left(c_{\mathrm{g}z}\mathscr{A}_w\right)=0\ , \tag{52}$$

where $\mathscr{A}_w = e_w\hat{\omega}^{-1}$ is the wave-action density.

e. Redimensionalization

The essential meso- and sub-mesoscale equations derived in section 2d are finally transformed back into the original coordinates and redimensionalized by applying the substitutions

$$\left(\mathbf{u}_0^{(0)},w_0^{(0)},b_0^{(0)},p_0^{(0)}\right) \rightarrow \left(\frac{\mathbf{U}}{U_m},\frac{W}{W_m},\frac{B}{B_m},\frac{P}{P_m}\right) \tag{53}$$

$$\left(\mathbf{u}_1^{(0)},w_1^{(0)},b_1^{(0)},p_1^{(0)}\right) \rightarrow \left(\frac{\tilde{\mathbf{u}}}{U_w},\frac{\tilde{w}}{W_w},\frac{\tilde{b}}{B_w},\frac{\tilde{p}}{P_w}\right) \tag{54}$$

$$\left(\mathbf{X}_{\mathrm{h}},Z,T\right) \rightarrow \left(\frac{\eta^2}{L_w}\mathbf{x}_{\mathrm{h}},\frac{\eta}{H_w}z,\frac{\eta}{T_w}t\right) \tag{55}$$

$$\left(f_0,N_0^2\right) \rightarrow \left(\frac{f}{f},a_w^2T_w^2N^2\right) \tag{56}$$

$$\left(\mathbf{k}_{\mathrm{h}}^{(0)},m,\hat{\omega}\right) \rightarrow \left\{L_w\mathbf{k}_{\mathrm{h}},H_wm,T_w\hat{\omega}\right)\ . \tag{57}$$

The dispersion and polarization relations (38) and (39) of the sub-mesoscale GWs now read

$$\hat{\omega} = \pm N\sqrt{\frac{|\mathbf{k}_\mathrm{h}|^2}{|\mathbf{k}_\mathrm{h}|^2\delta_{q0}+m^2}} \tag{58}$$

$$\{\tilde{w},\tilde{p},\tilde{\mathbf{u}}\} = \mathrm{i}\hat{\omega}\left\{\frac{1}{m},-\frac{\hat{\omega}}{|\mathbf{k}_\mathrm{h}|^2},-\frac{\mathbf{k}_\mathrm{h}}{|\mathbf{k}_\mathrm{h}|^2}\right\}\tilde{a}\,, \tag{59}$$

with $\tilde{a} = mN^{-2}\tilde{b}$, and their energy density is given by

$$\varepsilon_w = \frac{1}{4}\left(|\tilde{\mathbf{u}}|^2+|\tilde{w}|^2\delta_{q0}+\frac{1}{N^2}|\tilde{b}|^2\right) = \frac{1}{2N^2}|\tilde{b}|^2\,, \tag{60}$$

with the corresponding wave-action density $\mathscr{A} = \varepsilon_w\hat{\omega}^{-1}$.

The so-called ray equations, consisting of the eikonal equations (41) and (42) as well as the wave-action density equation (52), describe completely the sub-mesoscale dynamics, while the mesoscale dynamics is governed by the buoyancy, continuity and momentum equations (35) and (43) - (45). After back-transformation to the original coordinates and redimensionalization a coupled equation system for the interaction between meso- and sub-mesoscale GWs is obtained. Specifically, the mesoscale prognostic equations are

$$\nabla\cdot\mathbf{V} = 0 \tag{61}$$

$$\frac{\partial B}{\partial t}+N^2W = 0 \tag{62}$$

$$\frac{\partial P}{\partial z}-B = 0 \tag{63}$$

$$\frac{\partial\mathbf{U}}{\partial t}+f\mathbf{e}_z\times\mathbf{U}+\nabla_\mathrm{h}P = -\mathbf{k}_\mathrm{h}\frac{\partial}{\partial z}\left(c_{\mathrm{g}z}\mathscr{A}\right)\,, \tag{64}$$

where $\mathbf{V} = \mathbf{U}+W\mathbf{e}_z$. These equations are linear in the mesoscale variables since the corresponding wave amplitude is sufficiently low. All mesoscale-flow non-linearities disappear in the asymptotic

limit of small η. The sub-mesoscale-wave forcing acts via the convergence of the vertical flux of pseudomomentum $\mathbf{k}_\mathrm{h}\mathscr{A}$, as one finds from (44) and (51). The sub-mesoscale dynamics is given by

$$\left(\frac{\partial}{\partial t}+c_{\mathrm{g}z}\frac{\partial}{\partial z}\right)\omega = \mathbf{k}_\mathrm{h}\cdot\frac{\partial \mathbf{U}}{\partial t} \tag{65}$$

$$\left(\frac{\partial}{\partial t}+c_{\mathrm{g}z}\frac{\partial}{\partial z}\right)m = -\left\{\mathbf{k}_\mathrm{h}\cdot\frac{\partial \mathbf{U}}{\partial z}\pm\frac{\mathrm{d}N}{\mathrm{d}z}\sqrt{\frac{|\mathbf{k}_\mathrm{h}|^2}{|\mathbf{k}_\mathrm{h}|^2\delta_{q0}+m^2}}\right\} \tag{66}$$

$$\left(\frac{\partial}{\partial t}+c_{\mathrm{g}z}\frac{\partial}{\partial z}\right)\mathscr{A} = -\mathscr{A}\frac{\partial c_{\mathrm{g}z}}{\partial z}\,. \tag{67}$$

Vertical propagation of the waves is too fast for lateral propagation effects to matter within the time elapsing during the vertical propagation over a mesoscale vertical length scale. Note that (65) follows from (66) and the dispersion relation (58), and hence is not an independent equation in the sub-mesoscale model. The vertical group velocity $c_{\mathrm{g}z}$ may be calculated with the help of the current vertical wave number m.

Finally it is to be mentioned that the coupled equation system (61) - (67) is energy-conserving. For the mesoscale energy density $\varepsilon_m = 1/2\left(|\mathbf{U}|^2+B^2/N^2\right)$ one finds from (61) - (64) that its tendency obeys

$$\begin{aligned}\frac{\partial \varepsilon_m}{\partial t} &= \mathbf{U}\cdot\frac{\partial \mathbf{U}}{\partial t}+\frac{1}{N^2}B\frac{\partial B}{\partial t}\\ &= -\mathbf{U}\cdot\left\{f\mathbf{e}_z\times\mathbf{U}+\nabla_\mathrm{h}P+\mathbf{k}_\mathrm{h}\frac{\partial}{\partial z}\left(c_{\mathrm{g}z}\mathscr{A}\right)\right\}-\frac{1}{N^2}BN^2W\\ &= -\nabla\cdot(\mathbf{V}P)-\mathbf{U}\cdot\mathbf{k}_\mathrm{h}\frac{\partial}{\partial z}\left(c_{\mathrm{g}z}\mathscr{A}\right)\,,\end{aligned} \tag{68}$$

while it can be seen from the dimensional version of (48), (50) and (51) that the evolution of sub-mesoscale energy density $\varepsilon_w = \mathscr{A}\hat{\omega}$ is governed by

$$\frac{\partial \varepsilon_w}{\partial t} = -\frac{\partial}{\partial z}\left(\hat{\omega} c_{\mathrm{g}z} \mathscr{A}\right) - c_{\mathrm{g}z}\mathscr{A}\frac{\partial}{\partial z}\left(\mathbf{U}\cdot\mathbf{k}_{\mathrm{h}}\right) . \tag{69}$$

Therefore, the local total energy density $\varepsilon_t = \varepsilon_m + \varepsilon_w$ satisfies

$$\frac{\partial \varepsilon_t}{\partial t} = -\nabla\cdot(\mathbf{V}P) - \frac{\partial}{\partial z}\left(\omega c_{\mathrm{g}z}\mathscr{A}\right) . \tag{70}$$

As a result, if there is no mesoscale pressure flux at all boundaries, and no sub-mesoscale energy flux $\omega c_{\mathrm{g}z}\mathscr{A}$ at the vertical boundary of the domain, or if periodic boundary conditions hold, the spatially integrated total energy density E_t is conserved:

$$\frac{\mathrm{d}E_t}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int \varepsilon_t \, \mathrm{d}^3x = 0 . \tag{71}$$

3. Description of the numerical models

In this section the numerical code used for validation tests is described: The WKB code PincFloit-WKB is an implementation of the theory presented above. PincFloit without WKB sub-mesoscale-wave model, but instead in a setting explicitly resolving the sub-mesoscale waves is used for large-eddy simulations (LES), to provide data against which to validate the WKB model as well as its underlying theory. Since the GW dynamics is invariant with regard to rotation of the horizontal coordinate system, both codes have been used in 2D mode.

a. The PincFloit-WKB model

1) SUB-MESOSCALE FLOW: THE LAGRANGIAN WKB MODEL

The numerical implementation of the interaction between sub-mesoscale and mesoscale flow is achieved by coupling a Lagrangian phase-space ray tracer (Muraschko et al. 2015) to the mesoscale resolving model PincFloit. As can be read directly from the sub-mesoscale GW equations (66) - (67), along rays satisfying

$$\frac{\mathrm{d}z}{\mathrm{d}t} = c_{\mathrm{g}z} = \mp \frac{N|\mathbf{k}_\mathrm{h}|m}{(|\mathbf{k}_\mathrm{h}|^2 + m^2)^{\frac{3}{2}}} \tag{72}$$

vertical wave number and wave-action density develop according to

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \dot{m} = -\left\{ \mathbf{k}_\mathrm{h} \cdot \frac{\partial \mathbf{U}}{\partial z} \pm \frac{\mathrm{d}N}{\mathrm{d}z} \sqrt{\frac{|\mathbf{k}_\mathrm{h}|^2}{|\mathbf{k}_\mathrm{h}|^2 + m^2}} \right\} \tag{73}$$

$$\frac{\mathrm{d}\mathscr{A}}{\mathrm{d}t} = -\mathscr{A} \frac{\partial c_{\mathrm{g}z}}{\partial z} \,, \tag{74}$$

while the frequency can be obtained from the wave number and local mesoscale flow by the dimensional version of the full dispersion relation (40). In (73) the replacement $\delta_{q0} \to 1$ has been done, since in the case $q = 1$ one has $k^2 \ll m^2$ anyway. Direct implementation of this model leads, however, to serious numerical instabilities (Tabaei and Akylas 2007; Rieper et al. 2013a). These are due to caustic situations where rays cross, leading to multivalued wave numbers and wave-action densities. This can be avoided by taking a spectral perspective (Muraschko et al. 2015), where (72) and (73) indicate movement through a phase space spanned by vertical position and wave number. Wave-action density $\mathscr{A}(\mathbf{x},t)$ is then replaced by the spectral phase-space wave-action density $\mathscr{N}(\mathbf{x},m,t)$ (e.g. Dewar 1970; Olbers 1976; Bühler and McIntyre 1999; Hertzog

et al. 2002; Muraschko et al. 2015), developing along the rays according to

$$\frac{\mathrm{d}_\mathrm{r}\mathscr{N}}{\mathrm{d}t} = 0 \,, \tag{75}$$

with $\mathrm{d}_\mathrm{r}/\mathrm{d}t = \partial/\partial t + c_{\mathrm{g}z}\partial/\partial z + \dot{m}\partial/\partial m$ the phase-space material derivative. $\mathscr{A}$ can be retrieved from it by the wave number integral $\mathscr{A}(\mathbf{x},t) = \int dm\, \mathscr{N}(\mathbf{x},m,t)$ and the sub-mesoscale momentum flux, e.g., is obtained from

$$c_{\mathrm{g}z}\mathbf{k}_\mathrm{h}\mathscr{A} \rightarrow \int dm\, c_{\mathrm{g}z}\mathbf{k}_\mathrm{h}\mathscr{N} \,. \tag{76}$$

This requires reconstructing the full phase-space dependence of $\mathscr{N}$ from infinitely many rays. In a first discretization step one therefore collects rays carrying non-zero wave-action in a number of rectangular ray volumes with constant $\mathscr{N}$. In principle the individual ray velocities will deform these ray volumes arbitrarily strongly. In a second discretization step this deformation is simplified by prescribing the ray volumes to keep a rectangular shape. More details on this and the corresponding momentum-flux reconstruction are given by Muraschko et al. (2015) and Bölöni et al. (2016).

The sub-mesoscale momentum flux as well as energy density are smoothed after regridding over a window of 3×3 PincFloit finite-volume cell equivalents, see sections 3a.2) and 3a.3) below, in order to avoid artificial peaks resulting from sampling problems due to the ray discretization. Moreover, in the simulations the effect of sub-mesoscale horizontal group velocity is indeed small compared with that of the vertical group velocity, as found in theory, but still noticeable in the comparisons with the LES. Therefore, this effect has been incorporated by allowing the ray volumes

to also propagate in the horizontal direction for several simulations,

$$\frac{\mathrm{d_r}\mathbf{x}_\mathrm{h}}{\mathrm{d}t} = \mathbf{U}_\mathrm{h} \pm \frac{Nm^2}{(|\mathbf{k}_\mathrm{h}|^2 + m^2)^{\frac{3}{2}}} \frac{\mathbf{k}_\mathrm{h}}{|\mathbf{k}_\mathrm{h}|} = \mathbf{c}_\mathrm{gh} \ , \tag{77}$$

implying as well a generalized phase-space material derivative $\mathrm{d_r}/\mathrm{d}t = \partial/\partial t + \mathbf{c}_\mathrm{g} \cdot \nabla + \dot{m}\partial/\partial m$, with $\mathbf{c}_\mathrm{g}$ the 3D group velocity. Finally, the ray tracer has been supplemented by a simple saturation scheme (Bölöni et al. 2016) to account for turbulent wave breaking. The wave-action density of the sub-mesoscale GW packet is locally reduced, when its amplitudes reach the upper limit of static stability. Results show that the saturation is important for the total energy budget. In the present Boussinesq context, however, it has not contributed significantly to the instantaneous wave field distribution and simulation results discussed below.

2) MESOSCALE FLOW: PINCFLOIT

The pseudo-incompressible flow solver with implicit turbulence modeling (PincFloit), originally developed by Rieper et al. (2013b) to solve the pseudo-incompressible equations of Durran (1989), modified appropriately to integrate the Boussinesq equations (1) - (3), has been used at mesoscale resolution to simulate the resolved mesoscale flow. To account for the impact of the unresolved sub-mesoscale waves, the momentum equation has been supplemented by the corresponding convergence of horizontal pseudomomentum flux, as indicated by (64) and (76):

$$\frac{\mathrm{D}\mathbf{V}}{\mathrm{D}t} + f\mathbf{e}_z \times \mathbf{U} = -\nabla P + B\mathbf{e}_z - \frac{\partial}{\partial z} \int dm\, c_{\mathrm{g}z} \mathbf{k}_\mathrm{h} \mathcal{N} \ . \tag{78}$$

The latter is provided by the Lagrangian WKB code described above. As adumbrated by (78), in PincFloit the leading-order mesoscale dynamics is identified with the full resolved mesoscale

non-linear flow. Technically, PincFloit uses a finite-volume discretization with a staggered grid. Time integration is performed by an adaptive third-order Runge-Kutta scheme with a CFL criterion. Pressure is computed, using the non-divergence constraint (61), by solving the corresponding Poisson equation. The latter is done using a BiCGSTAB method (van der Vorst 1992). More details can be found in Rieper et al. (2013b).

3) THE COUPLED PINCFLOIT-WKB MODEL

PincFloit and the Lagrangian WKB model are coupled interactively, so as to simulate the transient interaction processes of resolved mesoscale, and unresolved sub-mesoscale GWs, as derived in section 2. At every Runge-Kutta sub-step, information is exchanged between the meso- and the sub-mesoscale dynamics. The Lagrangian WKB model determines the momentum flux convergence of the sub-mesoscale waves via the discretization of (76) and updates the mesoscale wind field which is then delivered to PincFloit. Hereafter, the latter integrates the Boussinesq equations (1) - (3) at mesoscale resolution. After that, the new wind and background values are provided to the Lagrangian WKB model, which solves the ray tracing equations (72), (73) and (75) (as well as (77), if intended), yielding an updated sub-mesoscale wave momentum flux and thus closing the circle.

Another remark is that this coupled system conserves the sum of mean-flow and wave energy too. From the Boussinesq equations (1) - (3) it can be derived for the mesoscale-flow energy density $\varepsilon_m = 1/2\left(|\mathbf{V}|^2 + B^2/N^2\right)$ that

$$\frac{\partial \varepsilon_m}{\partial t} = -\nabla\cdot\left(\mathbf{V}\left[\varepsilon_m + P\right]\right) - \frac{\partial}{\partial z}\left(\mathbf{U}\cdot\int c_{gz}\mathbf{k}_h \mathscr{N}\,\mathrm{d}m\right) + \frac{\partial \mathbf{U}}{\partial z}\cdot\int c_{gz}\mathbf{k}_h \mathscr{N}\,\mathrm{d}m\,, \tag{79}$$

while Bölöni et al. (2016) have shown for the sub-mesoscale wave energy density $\varepsilon_w = \int dm \hat{\omega} \mathcal{N}$ that

$$\frac{\partial \varepsilon_w}{\partial t} = -\frac{\partial}{\partial z} \int \hat{\omega} c_{\mathrm{g}z} \mathcal{N} \, \mathrm{d}m - \frac{\partial \mathbf{U}}{\partial z} \cdot \int c_{\mathrm{g}z} \mathbf{k}_{\mathrm{h}} \mathcal{N} \, \mathrm{d}m \, , \tag{80}$$

so that total energy $\int \varepsilon_t d^3x = \int (\varepsilon_m + \varepsilon_w) d^3x$ is conserved under suitable boundary conditions (zero or periodic) for the respective fluxes.

b. PincFloit-LES

In LES mode PincFloit is used to integrate the fully non-linear Boussinesq equations (1) - (3), with the above WKB sub-mesoscale-wave model switched off. Its resolution is chosen fine enough that the initial sub-mesoscale wave field is completely resolved, and that it captures wave-wave interactions, and interactions between all waves and the larger-scale turbulent eddies. Motivated by the results from corresponding benchmark tests (Remmler et al. 2015), small-scale turbulence is not parameterized by the implicit adaptive local deconvolution method (ALDM; see e.g. Hickel et al. (2006)), as originally implemented into PincFloit, but by a dynamic Smagorinsky method (Germano et al. 1991). The corresponding Smagorinsky coefficient is averaged over a local spatial window of 5×5 finite-volume cells so as to stabilize the scheme.

4. Numerical experiments

PincFloit-WKB and PincFloit-LES were used to simulate the propagation of a spatially confined wave packet in a uniformly stratified ($N = 0.02$ s^{-1}) atmosphere on an f-plane ($f = 10^{-4}$ s^{-1}, except in the test case COR, where $f = 0$) with zero initial ambient flow. It is well known from long-short-wave interaction theory (Tabaei and Akylas 2007; Van den Bremer and Sutherland 2014) that such a packet of small-scale waves is able to generate a mean flow consisting of mesoscale wave structures connected to a resonance phenomenon (see also section 4c). In turn, the mesoscale

waves may have an influence on the propagation of the wave packet. All simulations, investigating the resonant behaviour of various sub-mesoscale wave packets, are two-dimensional. The horizontal x-axis is chosen to point into the direction of $\mathbf{k}_\mathrm{h} = k\mathbf{e}_x$ and all initial fields are only dependent on x and z, as is then also the case henceforth. All models use periodic boundary conditions in x and z.

a. Initialization

Consequently, a locally monochromatic wave packet with horizontal and vertical wave numbers k and m_0, respectively, is initialized. It is vertically as well as horizontally confined, with a Gaussian-envelope amplitude $\hat{\mathscr{B}}$ characterized by the standard deviations σ_x and σ_z. Its buoyancy field is thus

$$b(x,z,t_0) = \hat{\mathscr{B}}(x,z)\cos(kx+m_0 z)\,, \tag{81}$$

$$\hat{\mathscr{B}}(x,z) = \frac{N^2}{m_0}\tilde{a}\exp\left[-\frac{(x-x_0)^2}{2\sigma_x^2}\right]\exp\left[-\frac{(z-z_0)^2}{2\sigma_z^2}\right]. \tag{82}$$

Herein the amplitude parameter $\tilde{a}$ is chosen so that at infinite σ_z static stability is given at $x = x_0$, i.e. $N^2 + \partial b/\partial z > 0$, for values $\tilde{a} < 1$. The initialization fields for the LES model are determined from the full rotational GW polarization relations (see e.g. Bühler (2009), §8.2)

$$\{\tilde{\mathbf{u}}, \tilde{w}\} = \left\{\mathrm{i}\frac{\mathbf{k}_\mathrm{h}\hat{\omega} - \mathrm{i}f\mathbf{e}_z\times\mathbf{k}_\mathrm{h}}{mN^2}\frac{\hat{\omega}^2 - N^2}{\hat{\omega}^2 - f^2}, \mathrm{i}\frac{\hat{\omega}}{N^2}\right\}\tilde{b}\,, \tag{83}$$

which lead, by the way, to the same initialization for a 2D wave packet with $l = 0$ as if one would use the leading-order polarization relations (59), except for the meridional wind component v,

yielding

$$\{u,w\}(x,z,t_0) = \left\{\frac{\hat{\omega}_0}{k}, -\frac{\hat{\omega}_0}{m_0}\right\} \hat{\mathscr{B}}(x,z)\,\sin(kx+m_0 z) \tag{84}$$

$$\{v,b\}(x,z,t_0) = \left\{-\frac{f}{k}, 1\right\} \hat{\mathscr{B}}(x,z)\,\cos(kx+m_0 z)\ , \tag{85}$$

where $\hat{\omega}_0$ denotes the intrinsic frequency of the initially monochromatic wave packet with wave numbers k and m_0, including rotational effects. In line with the hydrostatic sub-mesoscale-wave scaling regime (20) defined above, the choice falls on $\lambda_x = 2\pi/k = 1000$ m and $\lambda_z = -2\pi/m_0 = 100$ m, based on typical mesoscales like $2\,\sigma_x = L_m = 100$ km and $2\,\sigma_z = H_m = 1$ km (except for the case SCALE in section 4d). A typical mesoscale timescale for the wave packet test case is given by $T_m = f^{-1} = 10000$ s.

In PincFloit-WKB the initial fields of the wave packet are to be defined for the Lagrangian WKB model. Following Muraschko et al. (2015), a certain set of ray volumes, each carrying specific wave properties, is 'placed' in a rectangular 2.5 σ-environment around the initial center of the wave packet in both x- and z-direction, covering more than 98.7 % of wave energy density, as shown in figure 1. The initial phase-space wave-action density of the quasi-monochromatic wave packet is assumed to be

$$\mathscr{N}(x,z,m,t_0) = \begin{cases} \dfrac{\hat{\mathscr{B}}^2(x,z)}{2N^2\hat{\omega}_0}\dfrac{1}{\Delta m_0} & , m_{0,1} \le m \le m_{0,2} \\ 0 & , \text{else} \end{cases}\ , \tag{86}$$

where $m_{0,1} = m_0 - \Delta m_0/2$ and $m_{0,2} = m_0 + \Delta m_0/2$. It is then discretized by rectangles in physical space, and in wave-number space by two wave-number intervals each centered at wave numbers $m = m_{0,1} + \Delta m_0/4$ and $m = m_{0,2} - \Delta m_0/4$, constituting together phase-space ray volumes within

which $\mathscr{N}$ is taken to be constant. After initialization the ray volumes propagate through the phase space in accordance with the ray equations (72), (73) and (77), changing in response to the mesoscale flow physical location and wave number. Hence an initially quasi-monochromatic wave packet can develop quite complex spectra, as also discernible from figure 6 below.

Resolution is chosen as follows: In PincFloit-LES around 33 grid points in x- and 20 grid points in z-direction per initial sub-mesoscale wavelength are set. For PincFloit-WKB it turned out that 10 grid points per typical mesoscale length scale are sufficient in most of the test cases (see table 3). Parameters describing the general setup are listed in table 2.

b. Postprocessing

While PincFloit-WKB outputs directly the sub-mesoscale GW momentum flux, their energy density, and mesoscale wind and buoyancy fields, the output fields of PincFloit-LES contain both meso- and sub-mesoscale information, which have to be separated. Instead of a wave number filter as applied by Van den Bremer and Sutherland (2014), a running mean over two initial sub-mesoscale wavelengths in each spatial direction filters sub-mesoscale contributions from the mesoscale ones. The former are then obtained by subtracting the mesoscale part from the full field and subsequently, momentum flux and wave energy density are calculated via

$$\tilde{u}\tilde{w}(x_j,z_i,t_k) = \overline{\tilde{u}(x_j,z_i,t_k)\tilde{w}(x_j,z_i,t_k)} \tag{87}$$

$$\varepsilon_w(x_j,z_i,t_k) = \frac{1}{2}\overline{\left(|\tilde{\mathbf{u}}|^2(x_j,z_i,t_k)+\tilde{w}^2(x_j,z_i,t_k)+\frac{1}{N^2}\tilde{b}^2(x_j,z_i,t_k)\right)} \tag{88}$$

at each grid point (x_j,z_i) for each output time t_k, where the overbar indicates an additional running mean over two wavelengths. There is no significant sensitivity to the choice of the average interval

when comparing the results for an average over one up to three (and more) initial sub-mesoscale wavelengths.

c. Resonant interaction of mesoscale with sub-mesoscale GWs

It is known that energy exchange between long and short GWs is particularly strong, when a packet of small-scale GWs interacts with its induced large-scale mean flow under resonance conditions. Grimshaw (1977) found such resonant behaviour for a modulated wave train propagating along a horizontal channel, whereas Sutherland (2001), Tabaei and Akylas (2007) and Van den Bremer and Sutherland (2014) investigated wave-mean flow interactions for spatially localized wave packets propagating vertically through an unbounded stratified Boussinesq fluid. Tabaei and Akylas (2007), in particular, pointed out that flat wave packets, characterized by a stronger modulation in the vertical than in the horizontal, can lead to resonant forcing of large-scale waves. This resonance arises when the wave packet modulation scales are compatible with those of free, hydrostatic inertia GWs, that are natural mode solutions of the Boussinesq system. Furthermore, in the small-amplitude limit, where to leading order the wave packet envelope propagates vertically as a wave of permanent form, this resonance singles out inertia GWs whose vertical phase velocity matches the vertical group velocity of the wave packet. Tabaei and Akylas (2007) further speculated that this mechanism might be responsible for the generation of inertia GWs in the real atmosphere.

As noted in section 2b, the scaling regime considered by Tabaei and Akylas (2007) corresponds to the non-hydrostatic limit (19) which, under atmospheric conditions, exhibits a rather strong scale separation $L_w/L_m = O(10^{-4})$ and $H_w/H_m = O(10^{-2})$; as a result, the associated sub-mesoscale waves would be quite short: $H_w = L_w = O(10\ \mathrm{m})$. The hydrostatic regime, identified in (20), however, features sub-mesoscale GWs with length scales about one order of magnitude

smaller than typical grid point distances of nowadays global NWP models and of the same magnitude as the distances in regional limited area models (see section 4a). Consequently, in the latter such sub-mesoscale waves reside on the largest unresolved scale, and the long-short GW resonance of Tabaei and Akylas (2007) is relevant to this regime.

The resonance condition noted above for small-amplitude wave packets can be written in the present notation as

$$\hat{c}_{\mathrm{g}z}=\hat{c}_{\mathrm{p}zm}\ , \tag{89}$$

where $\hat{c}_{\mathrm{g}z}$ is the (intrinsic) vertical group velocity of the sub-mesoscale wave packet and $\hat{c}_{\mathrm{p}zm}$ the vertical phase velocity of the induced, mesoscale inertia GWs. Van den Bremer and Sutherland (2014) used this condition to estimate the phase line tilt of the generated large-scale waves in a non-rotational atmosphere. Here, based on (89), the geometry of the induced mesoscale inertia GW disturbance in the presence of rotation is discussed, following a kinematic approach analogous to that of the classical Kelvin ship wave pattern (see Whitham (1974), §12.4).

The resonance condition (89) combined with the dispersion relation (5) implies that the horizontal wave number k_m can be expressed in terms of the vertical wave number m_m of the induced mesoscale GWs (which are hydrostatic according to (4)):

$$k_m = \pm\frac{m_m}{N}\sqrt{\hat{c}_{\mathrm{g}z}^2 m_m^2 - f^2}\ . \tag{90}$$

As noted earlier, this assumes that: (i) The sub-mesoscale wave packet envelope, which acts as forcing, propagates as a wave of permanent form at constant speed $\hat{c}_{\mathrm{g}z}$; and (ii) the induced mesoscale waves are steady in the frame of the moving source. The validity of these assumptions will be tested numerically in section 4d.

In addition, one has generally $\partial m_m/\partial x = \partial k_m/\partial z$, which combined with (90) yields

$$\frac{\partial m_m}{\partial x} - \frac{\partial k_m}{\partial m_m}\frac{\partial m_m}{\partial z} = 0 \, . \tag{91}$$

This equation for m_m can be treated by the methods of characteristics. In the far field, all characteristics (straight lines) originate from the source (which appears as a point, $x = z = 0$, say, in the moving frame, with $z > 0$ behind the source), and it is found that

$$\frac{z}{x} = \mp\frac{1}{N}\frac{2\hat{c}_{gz}^2 m_m^2 - f^2}{\sqrt{\hat{c}_{gz}^2 m_m^2 - f^2}} \, . \tag{92}$$

Equation (92) determines m_m for given z/x, and the corresponding k_m follows from (90). It should be noted that $|z/x| \to \infty$ as $|m_m| \to f/\hat{c}_{gz}$; also, when $|m_m| \to \infty$, one has $|z/x| \approx 2|m_m|\hat{c}_{gz}/N \to \infty$. This suggests that there must be limiting characteristics $[z/x]_{\rm lim}$, where

$$\frac{\partial (z/x)}{\partial m_m} = 0 \, . \tag{93}$$

From this one obtains the critical wave number $|m_{m,\rm lim}| = (3/2f^2/\hat{c}_{gz}^2)^{1/2}$, which along with (92) gives $[z/x]_{\rm lim} = \mp 2\sqrt{2}\ f/N$. Thus, a Λ-like wave pattern behind the moving wave packet is expected as in the case of a ship wake. The opening angle α defined by the limiting characteristics is given by

$$\alpha = 2\,\tan^{-1}\left(\frac{1}{2\sqrt{2}}\frac{N}{f}\right) \, . \tag{94}$$

As can be seen from (19), (20), (24) and (29), in compressed coordinates the opening angle is $2\,\tan^{-1}(1/[2\sqrt{2}]) \approx 38.9^\circ$, independently of the scaling regime. Remarkably, this angle is identical to that of the classical Kelvin ship wave pattern on deep water (Whitham 1974).

Solving the characteristic equation (92) for m_m^2 gives

$$m_m^2 = \frac{1}{\hat{c}_{\mathrm{g}z}^2}\left\{\frac{1}{2}\left[f^2+\left(\frac{N}{2}\right)^2\left(\frac{z}{x}\right)^2\right] \pm \frac{N}{2}\left|\frac{z}{x}\right|\sqrt{\frac{1}{2}\left[\frac{1}{2}\left(\frac{N}{2}\right)^2\left(\frac{z}{x}\right)^2 - f^2\right]}\right\}. \tag{95}$$

Generally, according to (95) combined with (90), there are four different possible modes: two short-wave (corresponding to the '+' sign in (95)) and two long-wave ('−' sign) modes. The former are not relevant here, as their length scale is not in agreement with the mesoscale derived in our theory. Moreover, one of the two long-wave modes (where the phase contributions $k_m x$ and $m_m z$ have opposite signs), does not obey the radiation condition, so that wave energy would not be radiated away from the source. Thus, only one mode (where $k_m x$ and $m_m z$ have the same sign) is expected to prevail in the simulations (see figure 2), consistent also with figure 1 in Tabaei and Akylas (2007).

d. Test case studies

Theory and its numerical realization by the PincFloit-WKB model are validated in several test case studies: A reference simulation REF has been performed first, investigating the propagation of a relatively high-amplitude wave packet ($\tilde{a} = 0.5$) in a uniformly stratified atmosphere initially at rest. After that, the initial amplitude is varied to both lower and higher amplitudes (test cases AMPx) in order to demonstrate the robustness of the Lagrangian WKB model for a wide range of amplitudes of sub-mesoscale waves. Then a wave packet pushing the limits of the asymptotic scaling (see table 1) as well as the WKB (see section 2c) assumptions is initialized, thus establishing the wide applicability of the PincFloit-WKB model and the range of validity of the theory (test case SCALE). The test case COR examines the sensitivity on the ratio f/N by setting $f = 0$; strictly speaking, this regime is not represented by the theory as $\eta = 0$. Physically, it is to be understood

as a limit case for a tropical background. Last, the test case PSINC is used to compare LES simulations of the REF wave packet in a pseudo-incompressible atmosphere to the chosen Boussinesq dynamics. PSINC will be referred to in the discussion in section 5. Table 3 provides an overview of the essential test case-specific model setup and namelist parameters. According to (20), the chosen background parameters $N = 0.02\ \mathrm{s}^{-1}$ and $f = 10^{-4}\ \mathrm{s}^{-1}$ imply $\alpha = \eta \approx 0.07 = O(0.1)$.

Wave packets have been initialized in accordance with (82) and (84) - (86). Except for the SCALE test case, the spatial extent of the initialization box (roughly associated with the wave packet size) in PincFloit-WKB is 250 km $\times$ 2.5 km. The quantities shown in the following figures are in dimensional units. The spatial coordinates are given in kilometers, whereas the time axis is scaled by the inverse of the inertial frequency f. The 2D LES fields are visualized after mapping them to a 512×512 grid, with corresponding domain dimensions 500 km $\times$ 10 km - except for the case SCALE with 1000 km $\times$ 5 km.

Figure 3 shows the initial condition in the REF test case; the distribution of wave energy density (calculated via (88)) is displayed, where one can easily see that most of the sub-mesoscale energy of the wave packet is contained in a range of roughly $1.5\ X \times 1.5\ Z \mathrel{\hat{=}} 150$ km $\times$ 1.5 km, with the compressed coordinates X and Z from (29). Since the wave packet amplitude is $\tilde{a} = 0.5$ (and one may lose or gain factors as e.g. 2π in the scaling procedure), the wave packet energy density is somewhat smaller than theoretically assumed.

For the test case REF, three slightly different PincFloit-WKB model configurations are set up. First, a single-column ray tracer with 1D spatial ray propagation, i.e. only in the vertical, is initialized and no feedback of the resolved flow onto the sub-mesoscale wave packet is allowed (i.e. the right hand side of both (73) and (77) equals zero; shortcut: *PincFloit-WKB-1DNF*). Secondly, a single-column ray tracer is used which accounts for two-way scale interactions (i.e. only the right hand side of (77) equals zero; shortcut: *PincFloit-WKB-1D*). Note that, although

ray-volume propagation is 1D, the ray-volume amplitudes in $\mathscr{N}$ are not (see (86)), so that the sub-mesoscale wave packet induces a 2D mesoscale response. Thirdly, also horizontal ray volume propagation is allowed (shortcut: *PincFloit-WKB-1.5D*).

After a bit less than one inertial period ($t \approx 17.3$ h), one has a couple of findings. In the case of a non-energy-conserving system with no coupled interaction (figure 4(a)), one observes horizontally more prolonged mesoscale wave structures than in the validating LES simulation 4(d). The colored contours represent the wave packet-induced mesoscale horizontal winds. Furthermore, there is no wave packet deformation with PincFloit-WKB-1DNF, which is a result of the feedback process (to be compared with figure 4(b)). Horizontal propagation of the wave packet is rather weak compared to the vertical in terms of the characteristic mesoscales. Incorporating (77), with PincFloit-WKB-1.5D (figure 4(c)) a picture, which is qualitatively as well as quantitatively very close to the LES, is obtained.

The overlay of the theoretically derived wave structures of figure 2 on figure 4(a) shows that indeed, the resonance condition examined in section 4c is able to explain the lateral confinement of the induced waves: the mesoscale wave patterns are reminiscent of modes which develop in the wake of a ship, and the spatial extent of them is in good agreement to the area bounded by the limiting characteristics - except close to the wave packet, as it is no point source as assumed in the theoretical derivation. Interestingly, self-acceleration effects cause a lateral constriction of the induced structures close to the wave packet, so that the feedback-allowing simulations resemble the prediction even more with regard to the limiting characteristic. The vertical wavelengths are predicted by (95) to be somewhat longer than observed in these simulations and the phase line tilt is expected to be less pronounced; these small differences are explainable by the application of the far field approximation, the disregarded feedback from the induced waves and hence the requirement of a constant group velocity of the wave packet in section 4c, as underscored by the

longer wavelengths in figure 4(a). Nevertheless, the equations derived there seem to be a very good indicator for the prediction of the induced wave mode.

Furthermore, the induced structures are evocative of the ones induced by the small-amplitude non-hydrostatic wave packet of Tabaei and Akylas (2007): they resemble plain downward-propagating inertia GWs. Their period - extracted from the Hovmöller diagram in figure 5 - is very close to the inertial period $2\pi/f$.

It shall be stressed again that - in contrast to the assumption of Van den Bremer and Sutherland (2014) - within the first inertial period, the wave packet energy density distribution varies already significantly, as the group velocity undergoes partially strong and rapid changes: the peak values of energy density are found in the lower sector of the wave packet (not shown). When mesoscale momentum and vertical momentum shear build up, the vertical wave number - and due to (72) vertical group velocity too - of a part of the sub-mesoscale waves undergoes considerable changes according to (73), as found by investigating the time evolution of the discrete wave number spectrum (figure 6). The characteristics and phase lines of the wave mode predicted in section 4c might hence slightly differ from figure 2 or figure 4(a), respectively, when abandoning the requirement of a constant group velocity. Notably, in contrast to Tabaei and Akylas (2007) it can be reported that PincFloit-WKB is stable all the time although caustics are ubiquitous in physical space as can be seen in figure 6(b). Similar to Bölöni et al. (2016) it is also found (not shown) that the model conserves total energy very well (e.g. not more than 2% variation in REF at the end of the simulation time), as predicted in sections 2e and 3a.3).

The AMPx test cases substantiate the finding that the refractive feedback onto the wave packet becomes larger, the larger its (initial) amplitude is, as one can clearly see on the figures 7 and 8. In other words: as one moves towards small amplitudes, meso- and sub-mesoscales interact only weakly non-linearly and the wave packet propagation is very well approximated by its initial

group velocity. Consequently, the difference between the results of PincFloit-WKB-1DNF (figure 7(a)) and PincFloit-WKB-1D (figure 7(b)) is considerably smaller than in the REF test case.

Induced mesoscale momentum amplitudes, however, are about one order smaller in the small-amplitude case AMP1 compared to the REF simulations, well in proportion to the reduced momentum forcing by the sub-mesoscale wave packet. While the frequency of the induced mesoscale waves remains close to the inertial period independently of wave packet amplitude, their vertical extent evolving during the wave packet propagation becomes the larger, the less refraction appears, as in the case of a small-amplitude quasi-steady propagating disturbance, where the mesoscale wind is too weak to influence the wave packet.

In the case SCALE, a very flat wave packet of hydrostatic sub-mesoscale GWs is initialized, whose ratio of vertical to horizontal amplitude-variation length σ_z/σ_x is decreased by one order of magnitude, but whose total energy is kept constant compared to REF. Vertical amplitude variation is thus much stronger and horizontal amplitude variation weaker than originally assumed in theory. Beyond that, the vertical amplitude variation scale is now close to the vertical sub-mesoscale wavelength. Even though the horizontal scale separation is in accordance with the general regime, this is a case at the limit of the underlying theory.

Figure 9 shows again the wave packet energy density and the induced mesoscale horizontal momentum fields as in figure 4. Apart from the partly band-like representation of the wave packet due to a limited set of ray volumes - allowing a passable computing time - in PincFloit-WKB, which experiences strong wind shear changes, and a slight overestimation of the mesoscale amplitudes by PincFloit-WKB, the results from the parameterized model agree well with the LES model. An inertia GW is generated whose resonant amplification is broken off after approximately one inertial period simulation time due to wave packet energy spreading. The elongation of the induced mesoscale wave is hence smaller in the vertical than e.g. in AMP1 after that time, like in the

higher-amplitude cases REF or AMP2 after approximately two inertial periods, when the wave packet has experienced strong refraction effects.

The test case COR, where $f = 0$ in contrast to the case REF, investigates the sensitivity of the implementation in PincFloit-WKB on the value of f/N (figure 10). Energy is radiated stronger and further laterally than in the rotational REF case, which is reflected by the induced wind structures. This finding has been carved out in detail by Van den Bremer and Sutherland (2014) and Tabaei and Akylas (2007), and underscores again the wide applicability of PincFloit-WKB.

5. Résumé, discussion and outlook

The first question in our study was whether there are any sub-mesoscale GWs that can modify mesoscale dynamics at typical scales. This has been investigated within Boussinesq dynamics for typical mid-latitude tropospheric values of f/N, using scale analysis and comparing the magnitude of large-amplitude sub-mesoscale GW flux convergences with the self-consistent acceleration and heating in statically stable mesoscale GWs. A range of scales of high- and mid-frequency sub-mesoscale GWs has been identified where an impact on the mesoscale waves is possible. It encompasses the case of very small-scale high-frequency GWs investigated by Tabaei and Akylas (2007), but even more interesting appear hydrostatic GWs with scales that are just below the resolution of present-day limited-area numerical-weather-prediction codes. For these waves the vertical scale separation is $\eta = \sqrt{f/N}$, while the horizontal scale separation is η^2.

Using multi-scale asymptotics a large-amplitude WKB theory for the interaction between locally monochromatic sub-mesoscale GWs and a mesoscale flow has been derived. All nondimensional fields are expanded in terms of powers of η, and then the distinguished limit of small η is taken. This leads to separate but coupled equation systems, one describing explicitly resolved mesoscale dynamics, and the other depicting sub-mesoscale dynamics, consisting of verti-

cal ray equations for the sub-mesoscale wave properties. Direct coupling is established on the one hand by a modification of the mesoscale-momentum equation through the vertical convergence of sub-mesoscale pseudomomentum flux, and on the other hand by a change of sub-mesoscale vertical wave number and frequency by mesoscale wind-shear regions, or, though unaffected by sub-mesoscale fluxes, the vertical variation of background stratification. An interesting and useful result might seem to be that, other than in the case of the interaction between synoptic-scale flow and mesoscale inertia GWs, neither is horizontal propagation of the small-scale GWs a leading-order effect, nor are small-scale-GW horizontal flux convergences. Hence single-column approaches to sub-mesoscale-GW modeling in mesoscale resolving models appear well justified at first sight. This seems to be in contrast to previous findings about the relevance of lateral propagation of mesoscale waves by Senf and Achatz (2011), Sato et al. (2012) and Plougonven et al. (2017), suggesting certain limitations of single-column approaches. Indeed, caution is at place to not misinterpret results. To see this, note that the ratio between vertical and horizontal group velocity is $|c_{\mathrm{g}z}/|\mathbf{c}_{\mathrm{gh}}|| = ||\mathbf{k}_{\mathrm{h}}|/m| = H_w/L_w$. Hence the horizontal distance L covered by a wave packet propagating in the vertical over a distance H is $L = H\,L_w/H_w = H\sqrt{N_*/f}$, where the scalings (4) and (20) are used. The secondary importance of lateral propagation seen here is due to the fact that, if $H = H_m$ is taken then $L/L_m = H/L_m\sqrt{N_*/f} = \sqrt{f/N_*} \ll 1$, again using (4). Hence the sub-mesoscale wave packet travels over considerably less a horizontal distance than the scale characterizing mesoscale variations. On the other hand, the ultimate justification for single-column implementations would be that for all possible vertical distances coverable, at most the vertical model extent $H = H_{\mathrm{top}}$, the horizontal distance L covered should be less than a horizontal mesh distance Δx_{h}. This would imply $L/\Delta x_{\mathrm{h}} = H_{\mathrm{top}}/\Delta x_{\mathrm{h}}\sqrt{N_*/f} \ll 1$. Typically this condition cannot be met. It hence appears safer to take lateral propagation into account, and our simulations also show improved results if one does so.

The validity of the theory has been examined by its implementation into a mesoscale resolving Boussinesq model. The WKB fields of sub-mesoscale GW amplitudes and wave numbers are discretized and predicted by a Lagrangian ray tracer (Muraschko et al. 2015; Bölöni et al. 2016) that uses a spectral phase-space representation, thereby avoiding numerical instabilities due to caustics (Tabaei and Akylas 2007; Rieper et al. 2013a) and also allowing the potential development of a spectral sub-mesoscale GW field from initially locally monochromatic conditions. The approach is validated against simulations by a sub-mesoscale-resolving large-eddy code.

Test cases launching a 2D Gaussian wave packet of hydrostatic sub-mesoscale GWs furnish evidence of the applicability of the numerical approach and its underlying theory. A resonance effect occurs, which has been found in previous studies of Tabaei and Akylas (2007) and Van den Bremer and Sutherland (2014), and leads to the generation of mesoscale inertia GWs. Their characteristics correspond to free modes of the Boussinesq dynamics, which can be explained by a theoretical study of the resonance condition. The inertia GWs can have a strong impact onto sub-mesoscale wave packets by refraction, depending on wave amplitude, and eventually cause a saturation or breakup of the effective resonant energy transfer. Further case studies show that the approach, although designed for large-amplitude sub-mesoscale GWs, also works for low-amplitude sub-mesoscale GWs, and that it also performs reasonably well when the scale separation between mesoscale and sub-mesoscale is weaker than assumed in the theoretical derivations. Beyond that, the code seems to be usable for non-rotating cases as well.

The Boussinesq setting of our analysis does not say that non-Boussinesq effects are irrelevant. The vertical decrease of ambient density would play an important role in operational weather-forecast and climate models. This is left out here for the mere sake of simplicity, but corresponding generalizations as in Bölöni et al. (2016) seem straightforward. Moreover, in an analysis on the basis of fully compressible dynamics Achatz et al. (2017) identify compressible and elastic effects

in the interaction between near-inertial mesoscale waves and synoptic-scale flow. One of these are synoptic-scale pressure fluctuations that matter in a strongly stratified atmosphere. The other is an elastic mesoscale-wave term appearing in the synoptic-scale momentum equations. That term is most relevant for near-inertial waves but loses importance in the non-inertial frequency range. The most interesting sub-mesoscale waves, which are identified in the present study, are in the latter range. Hence, a fully compressible treatment seems a necessary extension of the present investigations based on Boussinesq theory; however, it is not clear, whether it will identify, at the scales of interest here, relevant non-Boussinesq effects, beyond those resulting from the ambient-density vertical dependence. For a first hint the REF case has been simulated with PincFloit-LES, but using it in the pseudo-incompressible mode (Durran 1989; Rieper et al. 2013b) instead of the Boussinesq mode (test case PSINC). Pseudo-incompressible dynamics captures the elastic effects arising in the analysis of Achatz et al. (2017) in the momentum equation. Figure 11 shows a snapshot of the simulation, to be compared to figure 4. Apparently at least in this case elastic dynamics does not seem to be of leading-order importance.

Our analysis shows that sub-mesoscale GWs can influence significantly a mesoscale flow, provided their amplitudes are large enough, and it also derives a theory and its numerical implementation for the efficient representation of such effects in mesoscale-resolving models. Whether sub-mesoscale GWs at the required scales and amplitudes are indeed present to a sufficient degree in the atmosphere, and by which processes they can be generated, certainly is another question that should be investigated in the future. Mesoscale wind-field spectra determined from aircraft data by Callies et al. (2014) and Bierdel et al. (2016), do not exhibit dissipation at the smallest observable scales so that one could imagine a continuation of these well into the sub-mesoscale range. A further indication is that GWs at very small scales seem to be a relevant issue in the stable planetary boundary layer (Sun et al. 2015). Should corresponding studies yield further support for

the relevance of sub-mesoscale GWs, an unavoidable next step would have to be extending the Boussinesq theory to an analysis of the compressible Euler equations.

Acknowledgments. U.A., R.K. and Ju.We. thank the German Research Foundation (DFG) for partial support through the research unit *Multiscale Dynamics of Gravity Waves (MS-GWaves)* and through Grants AC 71/8-1, AC 71/9-1, AC 71/10-1 and KL 611/25-1. U.A. and B.R. thank the German Federal Ministry of Education and Research (BMBF) for partial support through the program *Role of the Middle Atmosphere in Climate (ROMIC)* and through Grant 01LG1220A. B.R. acknowledges funding support from the CEA and the European project ARISE 2 (Horizon 2020, GAN653980). U.A., T.R.A. and Ja.Wi. thank *MISTI (MIT-Germany Seed Fund)* for making their collaboration possible.

APPENDIX

Derivation of a general scaling regime

As motivated at the beginning of section 2b, in order to identify a general scaling regime for the interaction between meso- and sub-mesoscale GWs, the sub-mesoscale flux convergence terms must be comparable to the leading mesoscale terms in the governing equations (1) - (3). Specifi-

cally, the relevant convergence terms scale as

$$
\begin{aligned}
-\nabla_{\mathrm{h}} \cdot (\mathbf{u}_w \mathbf{u}_w) &= O\left(\frac{U_w^2}{L_m}\right) && \text{(A1)}\\
-\frac{\partial}{\partial z}(\mathbf{u}_w w_w) &= O\left(\frac{U_w W_w}{H_m}\right) && \text{(A2)}\\
-\nabla_{\mathrm{h}} \cdot (\mathbf{u}_w w_w) &= O\left(\frac{U_w W_w}{L_m}\right) && \text{(A3)}\\
-\frac{\partial}{\partial z}(w_w w_w) &= O\left(\frac{W_w W_w}{H_m}\right) && \text{(A4)}\\
-\nabla_{\mathrm{h}} \cdot (\mathbf{u}_w b_w) &= O\left(\frac{f}{\Omega_w}\frac{U_w B_w}{L_m}\right) && \text{(A5)}\\
-\frac{\partial}{\partial z}(w_w b_w) &= O(0)\ , && \text{(A6)}
\end{aligned}
$$

where one has additionally made use of the GW polarization relations as given in (83); e.g. for the buoyancy-flux estimates: due to the phase shift between momentum and buoyancy, the fluxes obey the relations $\Re\{\tilde{\mathbf{u}}\tilde{b}^*\} = O(fL_w B_w) = O(f/\Omega_w\, U_w B_w)$, and $\Re\{\tilde{w}\tilde{b}^*\} = 0$, with tilde indicating the sub-mesoscale wave amplitudes and * denoting the complex conjugate. The convergence terms above are to be compared to the horizontal acceleration in the horizontal momentum equation, the buoyancy time derivative in the buoyancy equation, and the buoyancy in the vertical momentum equation for the mesoscale flow. These terms scale as

$$
\begin{aligned}
\frac{\partial \mathbf{u}_m}{\partial t} &= O\left(\frac{U_m}{T_m}\right) && \text{(A7)}\\
b_m &= O(B_m) && \text{(A8)}\\
\frac{\partial b_m}{\partial t} &= O\left(\frac{B_m}{T_m}\right)\ . && \text{(A9)}
\end{aligned}
$$

Using these scalings the following ratios for the horizontal momentum equation are found (always neglecting the modulationally unstable strongly non-hydrostatic sub-mesoscale regime in

(14)):

$$-\frac{\nabla_{\rm h}\cdot(\mathbf{u}_w\mathbf{u}_w)}{\partial\mathbf{u}_m/\partial t} = O\left(\frac{\eta}{a_w^2}\frac{\Omega_w^2}{N_*^2}\right) = \begin{cases} O\left(\dfrac{1}{\eta}\dfrac{L_w^2}{L_m^2}\right), & a_w \le \dfrac{f}{N_*} \\ O(\eta) \ll 1, & \dfrac{f}{N_*} < a_w \le 1 \end{cases} \tag{A10}$$

$$-\frac{\partial\left(\mathbf{u}_w w_w\right)/\partial z}{\partial\mathbf{u}_m/\partial t} = O\left(\frac{\eta}{a_w}\frac{\Omega_w^2}{fN_*}\right) = \begin{cases} O\left(\dfrac{L_w}{L_m}\right) \ll 1, & a_w \le \dfrac{f}{N_*} \\ O\left(\eta\dfrac{H_w}{L_w}\dfrac{N_*}{f}\right), & \dfrac{f}{N_*} < a_w \le 1 \end{cases} . \tag{A11}$$

The comparisons for the vertical momentum equation yield

$$\begin{aligned} -\frac{\nabla_{\rm h}\cdot(\mathbf{u}_w w_w)}{b_m} &= O\left(\frac{\eta}{a_w}\frac{f\Omega_w^2}{N_*^3}\right) \\ &= \begin{cases} O\left(\dfrac{L_w}{L_m}\dfrac{f^2}{N_*^2}\right) \ll 1, & a_w \le \dfrac{f}{N_*} \\ O\left(\eta a_w\dfrac{f}{N_*}\right) < O\left(\eta\dfrac{f}{N_*}\right) \ll 1, & \dfrac{f}{N_*} < a_w \le 1 \end{cases} \end{aligned} \tag{A12}$$

and

$$-\frac{\partial\left(w_w w_w\right)/\partial z}{b_m} = O\left(\eta\frac{\Omega_w^2}{N_*^2}\right) = \begin{cases} O\left(\eta\dfrac{f^2}{N_*^2}\right) \ll 1, & a_w \le \dfrac{f}{N_*} \\ O\left(\eta a_w^2\right) < O(\eta) \ll 1, & \dfrac{f}{N_*} < a_w \le 1 \end{cases} \tag{A13}$$

and for the buoyancy equation one obtains

$$-\frac{\nabla_{\rm h}\cdot(\mathbf{u}_w b_w)}{\partial b_m/\partial t} = O\left(\frac{L_w}{L_m}\right) \ll 1 . \tag{A14}$$

One sees that the sub-mesoscale wave field can impact the mesoscale flow only via the horizontal momentum equation. Two options exist: (i) The first is the impact of the low-frequency sub-mesoscale waves via horizontal momentum-flux convergence, see (A10). Equality be-

tween this flux term and the mesoscale-flow horizontal acceleration is reached when, using (6), $1 = (1/\eta)(L_w^2/L_m^2) = \eta^{2p-1}$. Hence in this regime $p = 1/2$. For an appreciable scale separation in the horizontal (see (6)), one would expect $\eta^{1/2} = O(10^{-1})$, say. This however implies very small η, so that the mesoscale-wave amplitude (9) that can be affected is very low. Without showing this in any detail, it is also mentioned that in this case the sub-mesoscale-wave frequency is dominated by the intrinsic part, while the Doppler term is small. This means that the mesoscale-flow impact on these waves is rather weak. (ii) The second and more interesting option involves the impact of the mid-frequency sub-mesoscale waves via vertical momentum-flux convergence, see (A11). Equality between flux convergence and the mesoscale-flow horizontal acceleration is reached here, using (4) and (6), when $1 = \eta a_w N_*/f = \eta^{2-p}$. Hence in this regime $p = 2$, and the horizontal length-scale separation η^2 and the vertical length-scale separation η can be small with $\eta = O(10^{-1})$, say, so that the mesoscale-wave amplitude that can be affected is stronger than in the first option. The possible range of η can be determined from the condition that the sub-mesoscale waves shall be in the mid-frequency range so that $f/N_* < a_w \leq 1$. Together with (7) and the requirement of a sufficiently strong scale separation $\eta \ll 1$ this implies (16).

References

Achatz, U., R. Klein, and F. Senf, 2010: Gravity waves, scale asymptotics and the pseudo-incompressible equations. *J. Fluid Mech.*, **663**, 120–147, doi:10.1017/S0022112010003411.

Achatz, U., B. Ribstein, F. Senf, and R. Klein, 2017: The interaction between synoptic-scale balanced flow and a finite-amplitude mesoscale wave field throughout all atmospheric layers: Weak and moderately strong stratification. *Q. J. R. Met. Soc.*, **143**, 342–361, doi:10.1002/qj.2926.

Alexander, M. J., and T. J. Dunkerton, 1999: A spectral parameterization of mean-flow forcing due to breaking gravity waves. *J. Atmos. Sci.*, **56**, 4167–4182.

Alexander, M. J., and Coauthors, 2010: Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Q. J. R. Meteorol. Soc.*, **136**, 1103–1124.

Andrews, D., and M. McIntyre, 1978a: An exact theory of nonlinear waves on a Lagrangian-mean flow. *J. Fluid Mech.*, **89**, 609– 646.

Andrews, D., and M. McIntyre, 1978b: On wave-action and its relatives. *J. Fluid Mech.*, **89**, 647–664.

Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584.

Bühler, O., 2009: *Waves and Mean Flows*. Cambridge Monographs on Mechanics, Cambridge University Press, 341 pp.

Bierdel, L., C. Snyder, S.-H. Park, and W. C. Skamarock, 2016: Accuracy of rotational and divergent kinetic energy spectra diagnosed from flight-track winds. *J. Atmos Sci.*, **73 (8)**, 3273–3286, doi:10.1175/JAS-D-16-0040.1.

Bölöni, G., B. Ribstein, J. Muraschko, C. Sgoff, J. Wei, and U. Achatz, 2016: The interaction between atmospheric gravity waves and large-scale flows: an efficient description beyond the non-acceleration paradigm. *J. Atmos. Sci.*, **73**, 4833–4852, doi:10.1175/JAS-D-16-0069.1.

Bretherton, F., 1969: On the mean motion induced by internal gravity waves. *J. Fluid Mech.*, **36**, 785–803.

Bretherton, F. P., 1966: The propagation of groups of internal gravity waves in a shear flow. *Quart. J. Roy. Met. Soc.*, **92**, 466–480.

Brune, S., and E. Becker, 2013: Indications of stratified turbulence in a mechanistic GCM. *J. Atmos. Sci.*, **70**, 231–247.

Bühler, O., and M. McIntyre, 1998: On non-dissipative wave-mean interactions in the atmosphere or oceans. *J. Fluid Mech.*, **354**, 301 – 343.

Bühler, O., and M. McIntyre, 1999: On shear-generated gravity waves that reach the mesosphere. Part II: Wave propagation. *J. Atmos. Sci.*, **56**, 3764–3773.

Bühler, O., and M. McIntyre, 2003: Remote recoil: a new wave-mean interaction effect. *J. Fluid Mech.*, **492**, 207–230.

Bühler, O., and M. McIntyre, 2005: Wave capture and wave-vortex duality. *J. Fluid Mech.*, **534**, 67–95, doi:10.1017/S0022112005004374.

Callies, J., R. Ferrari, and O. Bühler, 2014: Transition from geostrophic turbulence to inertia-gravity waves in the atmospheric energy spectrum. *Proc. Natl. Acad. Sci.*, **111 (48)**, 17 033–17 038.

Dewar, R. L., 1970: Interaction between hydromagnetic waves and a time-dependent, inhomogeneous medium. *Phys. Fluids*, **13 (11)**, 2710–2720, doi:10.1063/1.1692854.

Durran, D., 1989: Improving the anelastic approximation. *J. Atmos. Sci.*, **46**, 1453–1461.

Fritts, D., and M. Alexander, 2003: Gravity wave dynamics and effects in the middle atmosphere. *Rev. Geophys.*, **41 (1)**, 1003, doi:10.1029/2001RG000106.

Garcia, R. R., and S. Solomon, 1985: The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere. *J. Geophys. Res.*, **90**, 3850–3868.

Germano, M., U. Piomelli, P. Moin, and W. H. Cabot, 1991: A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids A*, **3**, 1760–1765.

Grimshaw, R., 1975: Nonlinear internal gravity waves in a rotating fluid. *J. Fluid Mech.*, **71**, 497–512, doi:10.1017/S0022112075002704.

Grimshaw, R. H. J., 1977: The modulation of an internal gravity-wave packet, and the resonance with the mean motion. *Stud. Appl. Math.*, **56 (3)**, 241–266, doi:10.1002/sapm1977563241.

Hansen, F., R. J. Greatbatch, G. Gollan, T. Jung, and A. Weisheimer, 2017: Remote control on NAO predictability via the stratosphere. *Q. J. R. Meteorol. Soc.*, **143 (1)**, 706–719, doi:10.1002/qj.2958.

Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Sheperd, and K. P. Shine, 1991: On the downward control of extratropical diabatic circulations by eddy-induced mean zonal forces. *J. Atmos. Sci.*, **48**, 651–678.

Hertzog, A., C. Souprayen, and A. Hauchecorne, 2002: Eikonal simulations for the formation and the maintenance of atmospheric gravity wave spectra. *J. Geophys. Res.*, **107**, doi:10.1029/2001JD000815.

Hickel, S., N. A. Adams, and J. A. Domaradzki, 2006: An adaptive local deconvolution method for implicit les. *J. Comp. Phys.*, **213 (1)**, 413 – 436, doi:10.1016/j.jcp.2005.08.017.

841 Holton, J. R., 1982: The role of gravity wave induced drag and diffusion in the momentum bud-
842 get of the mesosphere. *J. Atmos. Sci.*, **39**, 791–799, doi:10.1175/1520-0469(1982)039⟨0791:
843 TROGWI⟩2.0.CO;2.

844 Jia, L., and Coauthors, 2017: Seasonal Prediction Skill of Northern Extratropical Surface Temper-
845 ature Driven by the Stratosphere. *J. Climate*, doi:10.1175/JCLI-D-16-0475.1.

846 Joos, H., P. Spichtinger, and U. Lohmann, 2009: Orographic cirrus in a future climate. *Atmo-*
847 *spheric Chemistry and Physics*, **9**, 7825–7845, doi:10.5194/acp-9-7825-2009.

848 Kawatani, Y., K. Sato, T. J. Dunkerton, S. Watanabe, S. Miyahara, and M. Takahashi, 2010a: The
849 roles of equatorial trapped waves and internal inertia-gravitywaves in driving the quasi-biennial
850 oscillation. Part I: Zonal mean wave forcing. *J. Atmos. Sci.*, **67**, 963–980.

851 Kawatani, Y., K. Sato, T. J. Dunkerton, S. Watanabe, S. Miyahara, and M. Takahashi, 2010b: The
852 roles of equatorial trapped waves and internal inertia-gravitywaves in driving the quasi-biennial
853 oscillation. Part II: Three-dimensional distribution of wave forcing. *J. Atmos. Sci.*, **67**, 981–997.

854 Kawatani, Y., M. Takahashi, K. Sato, S. P. Alexander, and T. Tsuda, 2009: Global distribution of
855 atmospheric waves in the equatorial upper troposphere and lower stratosphere: AGCM simula-
856 tion of sources and propagation. *J. Geophys. Res.*, **114**, D01 102, doi:10.1029/2008JD010374.

857 Kidston, J., A. A. Scaife, S. C. Hardiman, D. M. Mitchell, N. Butchart, M. P. Baldwin, and
858 L. J. Gray, 2015: Stratospheric influence on tropospheric jet streams, storm tracks and surface
859 weather. *Nature Geosci.*, **8 (6)**, 433–440.

860 Kim, Y.-J., S. Eckermann, and H.-Y. Chun, 2003: An overview of the past, present and future
861 of gravity-wave drag parametrization for numerical climate and weather prediction models.
862 *Atmos.-Ocean*, **41**, 65–98.

Koch, S. E., and Coauthors, 2005: Turbulence and gravity waves within an upper-level front. *J. Atmos. Sci.*, **62**, 3885–3908.

Lindzen, R., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86**, 9707–9714, doi:10.1029/JC086iC10p09707.

Lott, F., and M. Miller, 1997: A new subgrid-scale orographic parameterization: its formulation and testing. *Q. J. R. Meteorol. Soc.*, **123**, 101–127.

McFarlane, N. A., 1987: The effect of orographically excited gravity wave drag on the circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.*, **44**, 1775–1800.

Muraschko, J., M. Fruman, U. Achatz, S. Hickel, and Y. Toledo, 2015: On the application of WKB theory for the simulation of the weakly nonlinear dynamics of gravity waves. *Quart. J. R. Met. Soc.*, **141**, 676–697, doi:10.1002/qj.2381.

Olbers, D. J., 1976: Nonlinear energy transfer and the energy balance of the internal wave field in the deep ocean. *J. Fluid Mech.*, **74 (2)**, 375–399, doi:10.1017/S0022112076001857.

Orr, A., P. Bechtold, J. Scinocca, M. Ern, and M. Janiskova, 2010: Improved middle atmosphere climate and forecasts in the ECMWF model through a nonorographic gravity wave drag parameterization. *Journal of Climate*, **23 (22)**, 5905–5926, doi:10.1175/2010JCLI3490.1, http://dx.doi.org/10.1175/2010JCLI3490.1.

Palmer, T. N., G. J. Shutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather-prediction models through an orographic gravity wave drag parametrization. *Q. J. R. Meteorol. Soc.*, **112**, 1001–1039.

Plougonven, R., A. de la Camara, V. Jewtoukoff, A. Hertzog, and F. Lott, 2017: On the relation between gravity waves and wind speed in the lower stratosphere over the southern ocean.

885 *J. Atmos. Sci*, **(0)**, in press, doi:10.1175/JAS-D-16-0096.1, URL http://dx.doi.org/10.1175/
886 JAS-D-16-0096.1, http://dx.doi.org/10.1175/JAS-D-16-0096.1.

887 Plougonven, R., and F. Zhang, 2014: Internal gravity waves from atmospheric jets and fronts. *Rev.*
888 *Geophys.*, **52**, 33–76, doi:10.1002/2012RG000419.

889 Remmler, S., S. Hickel, M. D. Fruman, and U. Achatz, 2015: Validation of large-eddy sim-
890 ulation methods for gravity wave breaking. *J. Atm. Sci.*, **72 (9)**, 3537–3562, doi:10.1175/
891 JAS-D-14-0321.1.

892 Ricard, D., C. Lac, S. Riette, R. Legrand, and A. Mary, 2013: Kinetic energy spectra characteristics
893 of two convection-permitting limited-area models arome and meso-nh. *Quart. J. R. Met. Soc.*,
894 **139 (674)**, 1327–1341, doi:10.1002/qj.2025.

895 Rieper, F., U. Achatz, and R. Klein, 2013a: Range of validity of an extended wkb theory for
896 atmospheric gravity waves: one-dimensional and two-dimensional case. *J. Fluid Mech.*, **729**,
897 330–363, doi:10.1017/jfm.2013.307.

898 Rieper, F., S. Hickel, and U. Achatz, 2013b: A conservative integration of the pseudo-
899 incompressible equations with implicit turbulence parameterization. *Mon. Wea. Rev.*, **141 (3)**,
900 861–886, doi:10.1175/MWR-D-12-00026.1.

901 Sato, K., S. Tateno, S. Watanabe, and Y. Kawatani, 2012: Gravity wave characteristics in the south-
902 ern hemisphere revealed by a high-resolution middle-atmosphere general circulation model. *J.*
903 *Atmos Sci.*, **69 (4)**, 1378–1396, doi:10.1175/JAS-D-11-0101.1, URL http://dx.doi.org/10.1175/
904 JAS-D-11-0101.1, http://dx.doi.org/10.1175/JAS-D-11-0101.1.

905 Scaife, A., and Coauthors, 2012: Climate change projections and stratosphere-troposphere inter-
906 action. *Clim. Dyn.*, **38 (9-10)**, 2089–2097, doi:10.1007/s00382-011-1080-7.

907 Scaife, A. A., J. R. Knight, G. K. Vallis, and C. K. Folland, 2005: A stratospheric influence
908 on the winter NAO and North Atlantic surface climate. *Geophys. Res. Lett.*, **32**, L18 715,
909 doi:10.1029/2005GL023226.

910 Scinocca, J. F., 2003: An Accurate Spectral Nonorographic Gravity Wave Drag Parameteriza-
911 tion for General Circulation Models. *J. Atmos. Sci*, **60**, 667–682, doi:10.1175/1520-0469(2003)
912 060⟨0667:AASNGW⟩2.0.CO;2.

913 Scinocca, J. F., and N. A. McFarlane, 2000: The parametrization of drag induced by stratified flow
914 over anisotropic orography. *Q. J. R. Meteorol. Soc.*, **126**, 2353–2393.

915 Senf, F., and U. Achatz, 2011: On the impact of middle-atmosphere thermal tides on the
916 propagation and dissipation of gravity waves. *J. Geophys. Res.*, **116**, D24 110, doi:10.1029/
917 2011JD015794.

918 Skamarock, W. C., 2004: Evaluating mesoscale nwp models using kinetic energy spectra. *Mon.*
919 *Wea. Rev.*, **132 (12)**, 3019–3032, doi:10.1175/MWR2830.1.

920 Sun, J., and Coauthors, 2015: Review of wave-turbulence interactions in the stable atmo-
921 spheric boundary layer. *Reviews of Geophysics*, **53 (3)**, 956–993, doi:10.1002/2015RG000487,
922 2015RG000487.

923 Sutherland, B. R., 2001: Finite-amplitude internal wavepacket dispersion and breaking. *J. Fluid*
924 *Mech.*, **429**, 343–380.

925 Tabaei, A., and T. R. Akylas, 2007: Resonant long-short wave interactions in an unbounded rotat-
926 ing stratified fluid. *Stud. Appl. Math.*, **119**, 271–296.

927 Van den Bremer, T. S., and B. R. Sutherland, 2014: The mean flow and long waves induced by
928 two-dimensional internal gravity wavepackets. *Phys. Fluids*, **26 (10)**, 106601, doi:10.1063/1.
929 4899262.

930 van der Vorst, H. A., 1992: Bi-cgstab: A fast and smoothly converging variant of bi-cg for the
931 solution of nonsymmetric linear systems. *SIAM J. Sci. Comput.*, **13 (2)**, 631–644, doi:10.1137/
932 0913035, https://doi.org/10.1137/0913035.

933 Warner, C. D., and M. E. McIntyre, 2001: An ultrasimple spectral parameterization for nonoro-
934 graphic gravity waves. *J. Atmos. Sci.*, **58**, 1837–1857.

935 Whitham, G. B., 1974: *Linear and Nonlinear Waves*. John Wiley & Sons, Inc., 636 pp.

936 Zhang, F., C. A. Davis, M. L. Kaplan, and S. E. Koch, 2001: Wavelet analysis and the governing
937 dynamics of a large-amplitude mesoscale gravity-wave event along the east coast of the united
938 states. *Q. J. Roy. Met. Soc.*, **127 (577)**, 2209–2245, doi:10.1002/qj.49712757702.

939 Zhang, F., S. E. Koch, and M. L. Kaplan, 2003: Numerical simulations of a large-
940 amplitude mesoscale gravity wave event. *Meteor. Atmos. Phys.*, **84 (3)**, 199–216, doi:10.1007/
941 s00703-002-0594-2.

LIST OF TABLES

960 TABLE 1. Overview of the appropriate scales for the interaction between mesoscale flow (subscript m) and
961 sub-mesoscale (subscript w) motions.

	Mesoscale GW	Sub-mesoscale GW
Vertical length scale	H_m	$H_w = \eta H_m$
Horizontal length scale	$L_m = \frac{N_*}{f} H_m$	$L_w = \eta^2 L_m$
Time scale	$T_m = f^{-1}$	$T_w = \eta T_m$
Scale of buoyancy amplitude	$B_m = \eta N_*^2 H_m$	$B_w = B_m$
Scale of horizontal wind amplitude	$U_m = \eta N_* H_m$	$U_w = U_m$
Scale of vertical wind amplitude	$W_m = \eta f H_m$	$W_w = \eta^{-1} W_m$
Scale of pressure fluctuation amplitude	$P_m = \eta N_*^2 H_m^2$	$P_w = \eta P_m$

TABLE 2. Synopsis of the relevant general model parameters. Here, n.a. indicates that the parameter is not used in this specific model setup. The time step Δt is determined through the CFL criterion; due to stability reasons there is an upper threshold for Δt in PincFloit-WKB. There the total number of ray volumes corresponds to the product of the number of grid cells in the 5 σ-box, and the corresponding number per grid cell and spatial direction $\tilde{n}_{x,\mathrm{ray}}$ or $\tilde{n}_{z,\mathrm{ray}}$, respectively. The smoothing parameter indicates the total number of grid cells in both directions, which are used for the respective local smoothing. A value of 3 means e.g. a smoothing over one gridcell and its direct neighbours (3×3 - window).

Configuration parameter	PincFloit-WKB	PincFloit-LES
Boundary conditions	periodic in x and z*	
Resolution $\Delta x \times \Delta z$	*test case-dependent*	30.5 m $\times$ 4.9 m
Time step Δt	dynamic, CFL $= 0.5, \Delta t_{\max} = 2$ s	dynamic, CFL $= 0.5, \Delta t_{\max} = 1$ s
Wave number interval Δm_0	$\Delta m_0 = 10^{-4}$ m^{-1}	n.a.
Number of ray volumes n_{ray}	$n_{\mathrm{ray}} = \frac{5\,\sigma_x}{\Delta x}\tilde{n}_{x,\mathrm{ray}} \times \frac{5\,\sigma_z}{\Delta z}\tilde{n}_{z,\mathrm{ray}}$	n.a.
Smoothing parameter	Momentum flux: $n_{\mathrm{s,uw}} = 3$	Dynamic Smagorinsky: $n_{\mathrm{s,dyn}} = 5$

* Except for the pseudo-incompressible simulation PSINC, where rigid boundaries are assumed in the vertical.

TABLE 3. Essential parameters for the initialization of a sub-mesoscale wave packet in different test cases. Here, n.a. indicates that the parameter is not used in this specific test case setup, i.e. there is no PincFloit-WKB simulation for the case PSINC. The frequency $\hat{\omega}_0 = \hat{\Omega}_+$ shall indicate that the positive branch of the dispersion relation (58) is chosen. Except for the case COR it is $f = 10^{-4}$ s^{-1}.

Test case	REF	AMP1	AMP2	COR	PSINC	SCALE
Simulation time t	$12\ f^{-1} = 120000$ s $= 2000$ min ≈ 33.3 h					
Model domain $L_x \times L_z$	500 km $\times$ 10 km $\Leftrightarrow 5\ L_m \times 10\ H_m$					$10\ L_m \times 5\ H_m$*
GP number P.F.-LES $n_x \times n_z$	16384×2048					32768×1024
GP number P.F.-WKB $n_x \times n_z$	51×101				n.a.	101×401
Resolution P.F.-WKB $\Delta x \times \Delta z$	10 km $\times$ 0.1 km				n.a.	10 km $\times$ 0.025 km
Ray vol's P.F.-WKB $\tilde{n}_{x,\mathrm{ray}} \times \tilde{n}_{z,\mathrm{ray}}$	20×5				n.a.	20×5
Amplitude $\tilde{a}$	0.5	0.1	1.0	0.5	0.5	0.5
Horizontal wave number k	$2\pi/1000$ m^{-1}					
Vertical wave number m_0	$-2\pi/100$ m^{-1}					
Frequency $\hat{\omega}_0$	$\hat{\Omega}_+$					
Initial vertical position z_0	2 km					
Initial horizontal position x_0	250 km					500 km
Vertical standard deviation σ_z	0.5 km					0.158 km
Horizontal standard deviation σ_x	50 km					158.113 km
Background stratification N	0.02 s^{-1}					
Coriolis parameter f	10^{-4} s^{-1}			0		10^{-4} s^{-1}

* In PincFloit-WKB $10\ L_m \times 10\ H_m$ are used.

LIST OF FIGURES

according to: (a) PincFloit-WKB-1.5D; (b) PincFloit-LES. Note the different shown extent of the domain compared to the previous figures; though, the vertical simulation domain in PincFloit-WKB-1.5D has been 10 km, so that ray volumes do not reappear at the model bottom in (a). 71

Fig. 10. Test case COR with initial amplitude $\tilde{a} = 0.5$ and $f = 0$ at $t \approx 17.3$ h. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in $\mathrm{m}^2\ \mathrm{s}^{-2}$, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in $\mathrm{m\ s}^{-1}$, colored contours), according to: (a) PincFloit-WKB-1.5D; (b) PincFloit-LES. Note the different color range compared to figure 4. 72

Fig. 11. Test case PSINC in a pseudo-incompressible background with initial amplitude $\tilde{a} = 0.5$ at $t \approx 17.3\ \mathrm{h} \approx 2\pi/f$. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in $\mathrm{m}^2\ \mathrm{s}^{-2}$, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in $\mathrm{m\ s}^{-1}$, colored contours), according to PincFloit-LES. This figure is to be compared with figure 4. 73

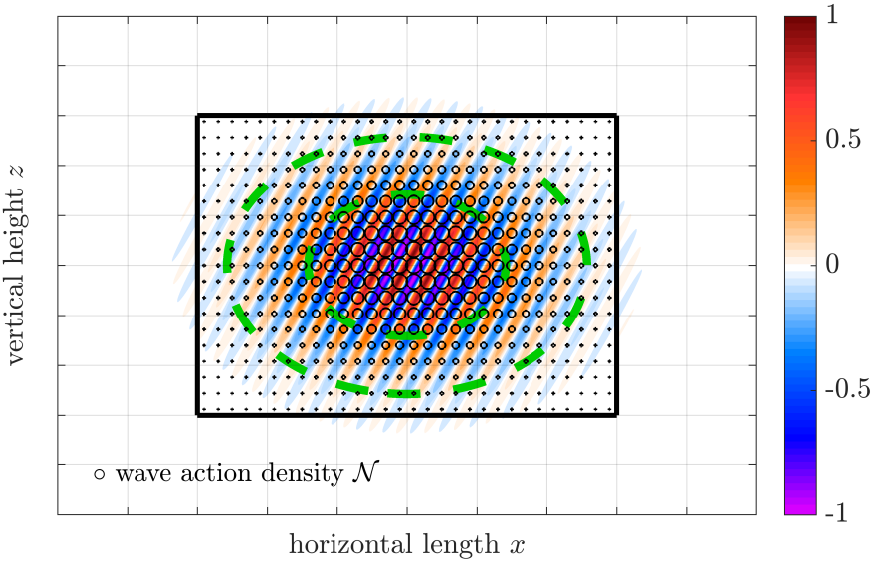


FIG. 1. Schematic illustration of the initial coverage of the wave packet in physical space with a discrete set of ray volumes, for simplicity sketched as circles. The wave structures represent e.g. the initial buoyancy field. The peak amplitude is chosen to be $\tilde{a} = 1$. The green dashed lines mark amplitude contours with $\tilde{a} = 0.5$ and $\tilde{a} = 0.1$. The size of the circles mirrors the corresponding modulus of the phase-space wave-action density $\mathcal{N}$. The box covers an area of 5 $\sigma_x \times 5$ σ_z.

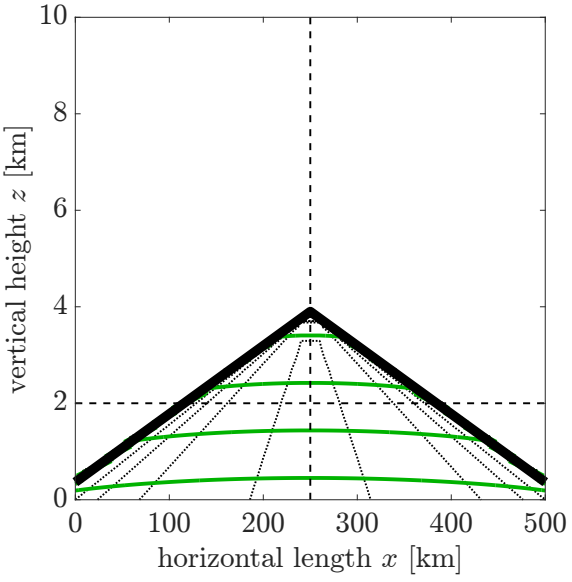
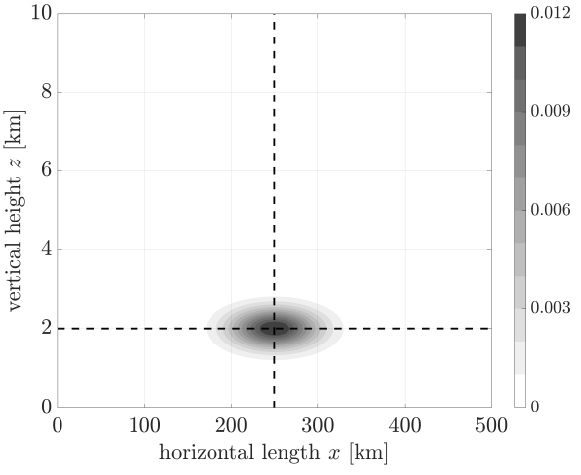
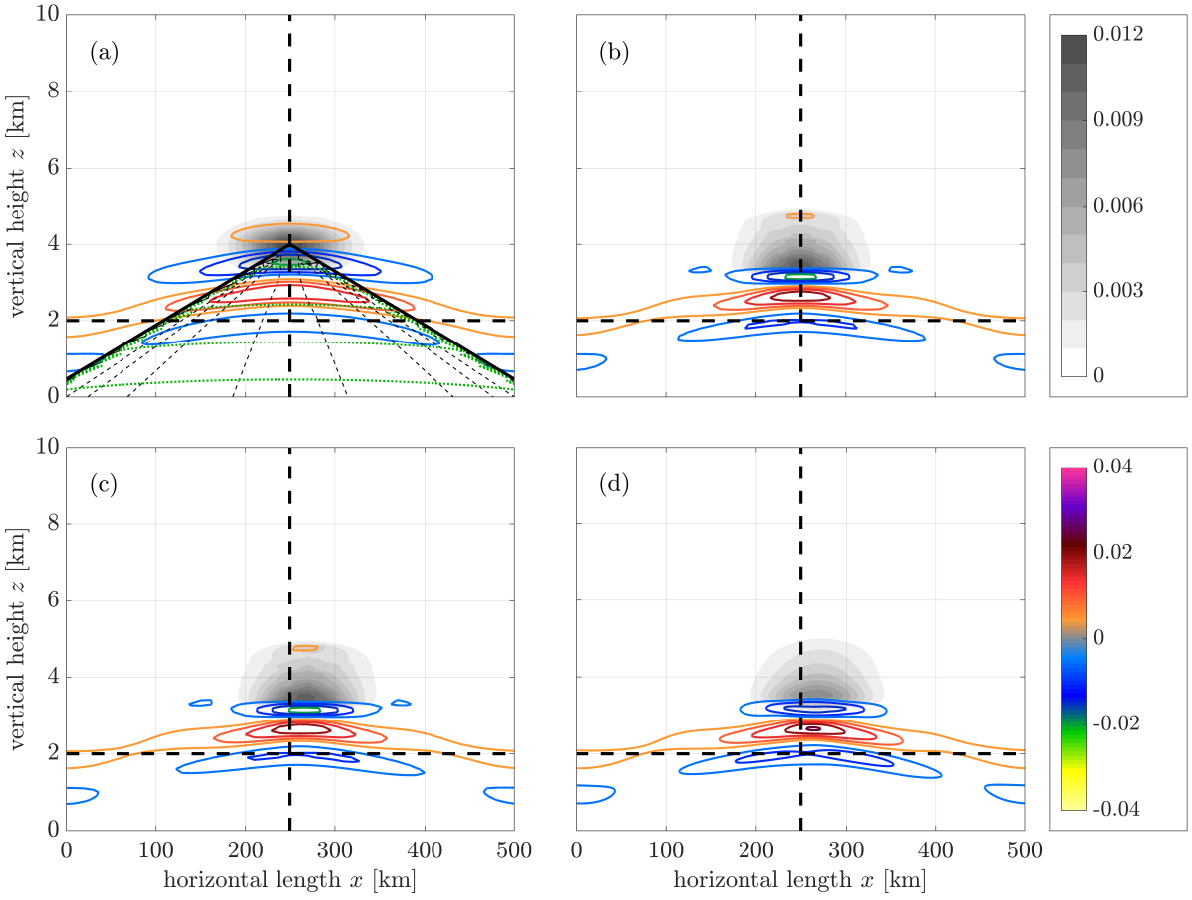


FIG. 2. Induced mesoscale wave pattern according to the kinematic wave theory of section 4c. The limiting characteristics $[z/x]_{\text{lim}}$ (black, solid line), lines of constant wave number $|\mathbf{k}_m|$ (dotted lines) and lines of constant phase (green lines) for a sub-mesoscale wave packet as in section 4d, case REF or AMPx, are shown at a point of time when the center of the wave packet is located at $(250, 3.9)$ km. The initial position of the center of the wave packet chosen for the test case studies is indicated by the intersection of the dashed lines.



1041 FIG. 3. Initial condition in the test case REF: Spatial distribution of energy density of the sub-mesoscale wave
1042 packet ε_w (in $\mathrm{m}^2\ \mathrm{s}^{-2}$) in PincFloit-LES.



1043 FIG. 4. Test case REF with initial amplitude $\tilde{a} = 0.5$ at $t \approx 17.3$ h $\approx 2\pi/f$. Spatial distribution of energy
1044 density of the sub-mesoscale wave packet ε_w (in m^2 s^{-2}, gray shades), and of the induced mesoscale hori-
1045 zontal wind speed in x-direction U (in m s^{-1}, colored contours), according to: (a) PincFloit-WKB-1DNF; (b)
1046 PincFloit-WKB-1D; (c) PincFloit-WKB-1.5D; (d) PincFloit-LES. In addition, (a) is overlayed with the theoret-
1047 ically derived phase lines (phase difference: π) and characteristics from figure 2.

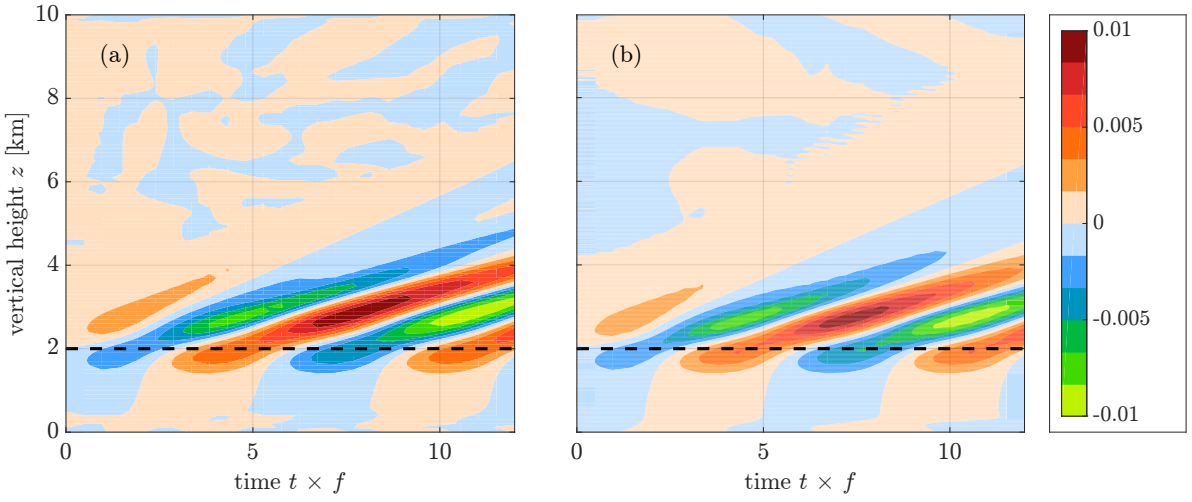


FIG. 5. Test case REF with initial amplitude $\tilde{a} = 0.5$. Hovmöller diagram of the horizontally averaged, induced mesoscale horizontal wind speed in x-direction U (in m s^{-1}), according to: (a) PincFloit-WKB-1.5D; (b) PincFloit-LES.

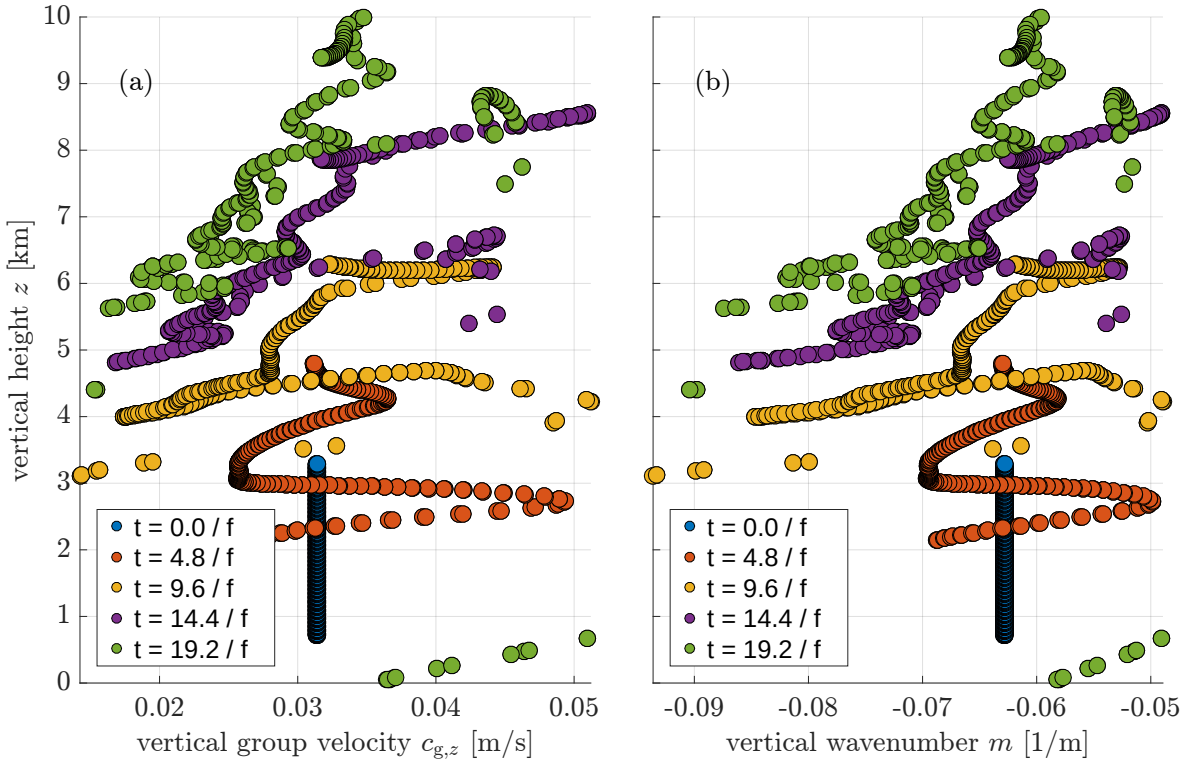


FIG. 6. Test case REF with initial amplitude $\tilde{a} = 0.5$. (a) Vertical distribution of the vertical group velocity c_{gz}; (b) vertical distribution of the vertical wave number m of the ray volumes in a column close to the horizontal center of the wave packet according to PincFloit-WKB-1D. The simulation time has been extended for this specific run to $t = 19.2/f \approx 53.3$ h. Different colors indicate different time instances. Due to the periodic boundary conditions, one observes ray volumes propagating through the model top and reappearing at the model bottom close to the end of the simulation.

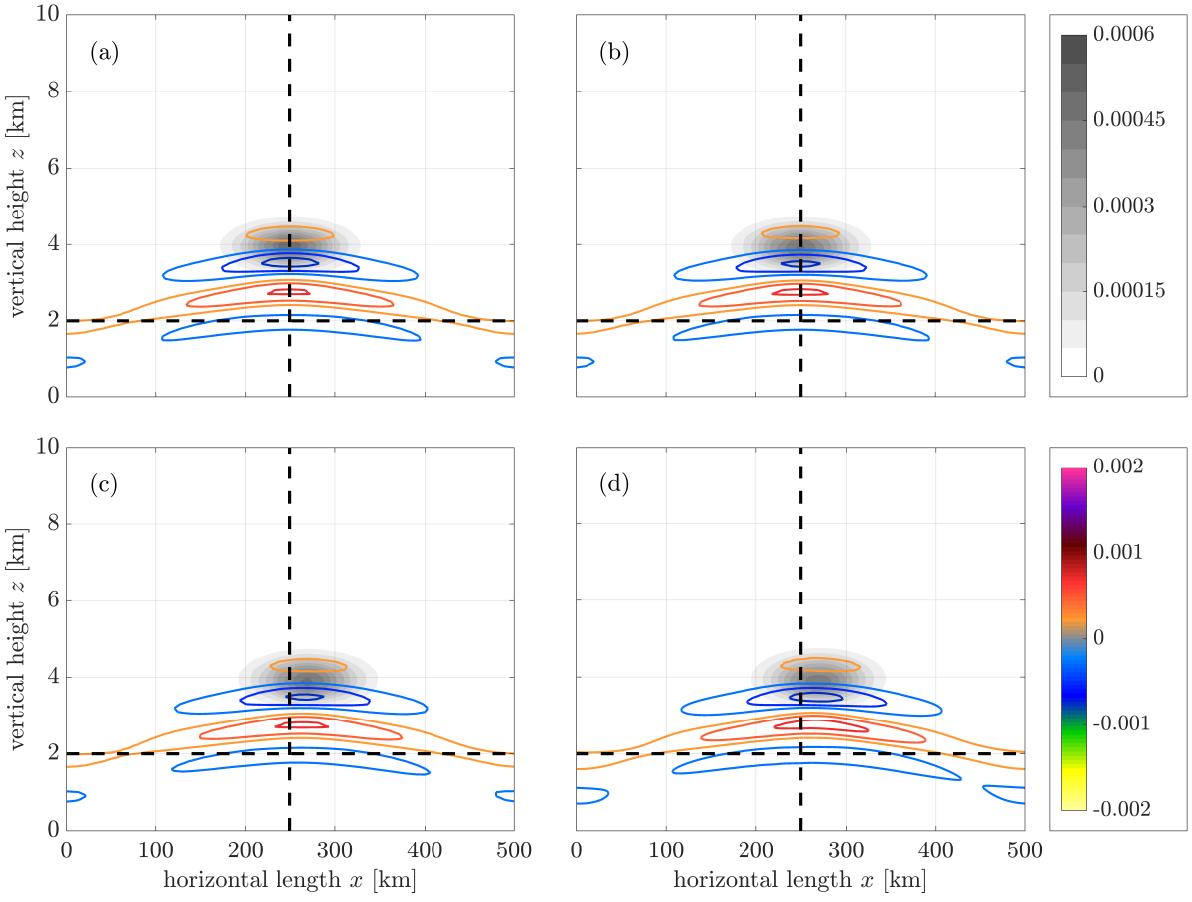


FIG. 7. Test case AMP1 with initial amplitude $\tilde{a} = 0.1$ at $t \approx 17.3$ h $\approx 2\pi/f$. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in m^2 s^{-2}, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in m s^{-1}, colored contours), according to: (a) PincFloit-WKB-1DNF; (b) PincFloit-WKB-1D; (c) PincFloit-WKB-1.5D; (d) PincFloit-LES. Note the different color and gray shading range compared to the previous figures.

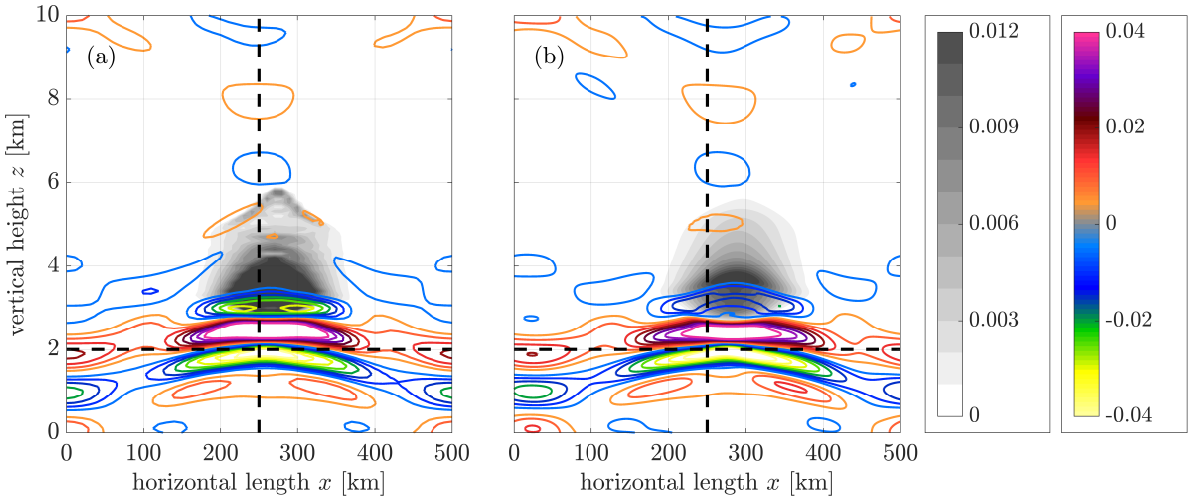


FIG. 8. Test case AMP2 with initial amplitude $\tilde{a} = 1.0$ at $t \approx 17.3$ h $\approx 2\pi/f$. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in m^2 s^{-2}, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in m s^{-1}, colored contours), according to: (a) PincFloit-WKB-1.5D; (b) PincFloit-LES.

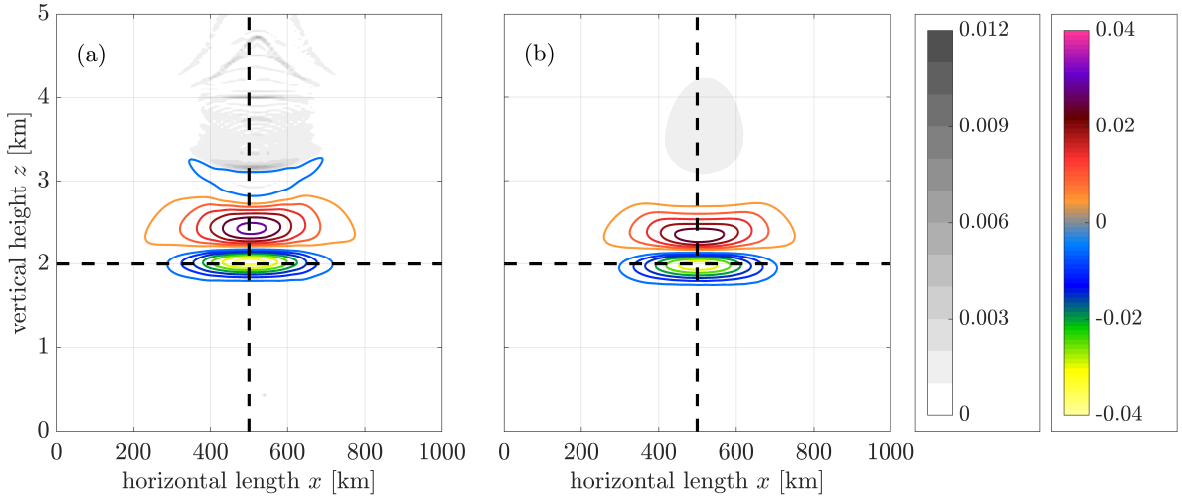


FIG. 9. Test case SCALE with initial amplitude $\tilde{a} = 0.5$ at $t \approx 17.3$ h $\approx 2\pi/f$. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in m^2 s^{-2}, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in m s^{-1}, colored contours), according to: (a) PincFloit-WKB-1.5D; (b) PincFloit-LES. Note the different shown extent of the domain compared to the previous figures; though, the vertical simulation domain in PincFloit-WKB-1.5D has been 10 km, so that ray volumes do not reappear at the model bottom in (a).

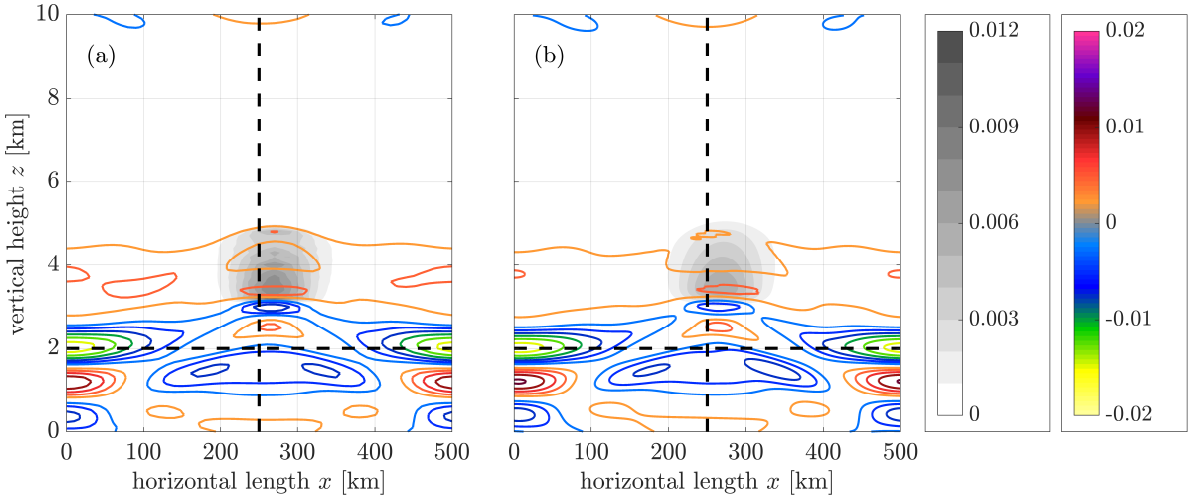


FIG. 10. Test case COR with initial amplitude $\tilde{a} = 0.5$ and $f = 0$ at $t \approx 17.3$ h. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in m^2 s^{-2}, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in m s^{-1}, colored contours), according to: (a) PincFloit-WKB-1.5D; (b) PincFloit-LES. Note the different color range compared to figure 4.

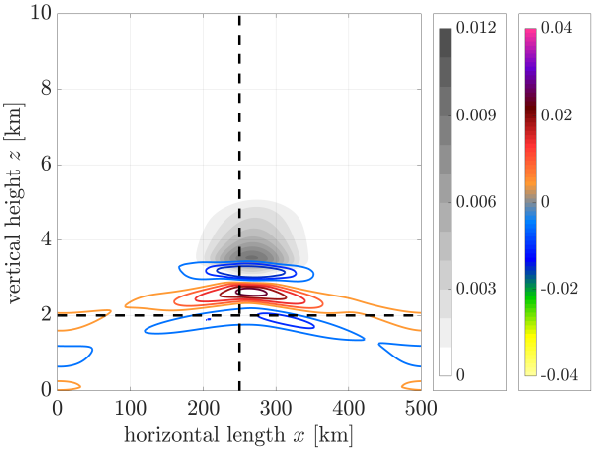


FIG. 11. Test case PSINC in a pseudo-incompressible background with initial amplitude $\tilde{a} = 0.5$ at $t \approx 17.3\ \mathrm{h} \approx 2\pi/f$. Spatial distribution of energy density of the sub-mesoscale wave packet ε_w (in $\mathrm{m}^2\ \mathrm{s}^{-2}$, gray shades), and of the induced mesoscale horizontal wind speed in x-direction U (in $\mathrm{m\ s}^{-1}$, colored contours), according to PincFloit-LES. This figure is to be compared with figure 4.