



Three-layer stratified exchange flows: hydraulically controlled transition to turbulence

Amir Atoufi¹ , Lu Zhu¹ , Adrien Lefauve^{1,2,3} , John R. Taylor¹ ,
Rich R. Kerswell¹ , Stuart B. Dalziel¹ , Gregory A. Lawrence⁴  and
Paul F. Linden¹ 

¹Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

²Grantham Institute - Climate Change and the Environment, Imperial College London, SW7 2AZ, UK

³Department of Civil and Environmental Engineering, Imperial College London, SW7 2BU, UK

⁴Department of Civil Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Corresponding author: Lu Zhu, lz447@cam.ac.uk

(Received 11 November 2024; revised 30 April 2025; accepted 10 June 2025)

Buoyancy-driven exchange flows in geophysical contexts often exhibit significant interfacial turbulence leading to a partially mixed intermediate layer between two counterflowing layers. In this paper we perform a three-layer hydraulic analysis of such flows, highlighting the dynamical importance of the middle mixed layer. Our analysis is based on the viscous, shallow water, Boussinesq equations and includes the effects of mixing as a non-hydrostatic pressure forcing. We demonstrate the superior predictive accuracy of three-layer hydraulics over the more classical two-layer approach by applying it to direct numerical simulation data in stratified inclined duct exchange flows where turbulence is controlled by a modest slope of the duct. The three-layer model predicts a region bounded by two control points in the middle of the duct, linked to the onset of instability and turbulence, whereas a two-layer model only predicts one control point. We show that the nonlinear characteristics of the three-layer model correspond to linear long waves perturbing a three-layer mean flow. We also provide the first evidence of long-wave resonance, as well as resonance between long and short waves, and their connection to turbulence. These results challenge current parameterisations for turbulent transport, which typically overlook long waves and internal hydraulics induced by streamwise variations of the flow.

Key words: shallow water flows, stratified flows, transition to turbulence

1. Introduction

Stratified exchange flows occur whenever two regions containing fluids with different densities are connected by a narrow duct. These flows have many applications in engineering (e.g. building ventilation, ducts and tunnels), as well as geophysics (e.g. between ocean basins through deep-sea conduits, estuaries and straits). For instance, in estuarine environments the intrusion of salty seawater directly affects coastal habitats. The measurement of turbulence and vertical mixing is challenging due to the high turbulent fluxes (Geyer, Scully & Ralston 2008) and spatiotemporal variability of the inflows and outflows (Burchard *et al.* 2019). Such variability in the exchange flow is partially controlled by the channel morphology, which ties turbulent mixing to internal hydraulics. For instance, Chant & Wilson (2000) showed that velocity and density have significant along-channel variability in the Hudson River estuary. Similarly, mixing in straits, for example, in the Straits of Gibraltar and Bab el Mandab, is often controlled by internal hydraulic processes (Farmer & Armi 1988; Pratt *et al.* 1999).

The internal hydraulics of exchange flows have mostly been studied using two-layer models (Armi 1986; Lawrence 1990; Dalziel 1991). However, when there is significant mixing between the layers, a third, middle layer of intermediate density is formed. Such a three-layer flow was studied in the Strait of Gibraltar by Bray, Ochoa & Kinder (1995), who showed with hydrographic observations that the interface was not a sharp boundary but a thick, stratified interfacial layer carrying a substantial portion of the total momentum. The study highlighted the significance of this middle layer due to its shear and demonstrated that two-layer hydraulic models are inadequate for predicting momentum fluxes. Similar three-layer configurations have also been observed in the Strait of Bab al Mandab (Smeed 2000) and the Dardanelles Strait connecting the Aegean Sea to the Sea of Marmara (Jarosz *et al.* 2012). In these situations, two-layer hydraulic models do not accurately predict hydraulic controls (i.e. state of maximal exchange flow rate), and more complex multi-layer models are needed (Sannino, Carillo & Artale 2007). The hydraulic control of multi-layer flows was first discussed by Benton (1954) from an energy perspective. Baines (1988) discussed the effects of upstream changes in the hydraulic properties of stratified multi-layer flow over an obstacle and showed that depending on the obstacle height, a hydraulic jump or time-dependent rarefaction might occur. Lane-Serff, Smeed & Postlethwaite (2000) used the 'hydraulic functional' (Gill 1977; Dalziel 1991) to study multi-layer exchange flows and showed good agreement between a three-layer model and a lock-exchange experiment.

Despite these studies, significant questions remain unresolved in three-layer flows. For example, Pratt *et al.* (1999) assessed three-layer properties of flow in Bab el Mandab based on acoustic Doppler current profiler velocity measurements. They identified subcritical or supercritical hydraulic flow regimes based on interfacial wave pairs propagating in the same or opposite directions with respect to a given mode. However, the mechanism behind the propagation of these interfacial waves and the local mixing rate is not yet adequately understood.

The fundamental difference between two- and three-layer internal hydraulics of stratified exchange flows is the possibility of strong interactions between interfacial waves in the presence of the two interfaces separating the three layers. In idealised stratified shear flow and infinitely long domains, the resonant interactions of short-wave instabilities have been studied extensively and understood (Smyth & Carpenter 2019). However, none of the previous studies have explored the possibility of long-wave resonant interactions between interfaces and their potential impact on the evolution of the flow and the triggering of instabilities. Moreover, in realistic conditions, the domain length prescribes finite wavelengths for the longest waves. The resonant interaction between long and short waves

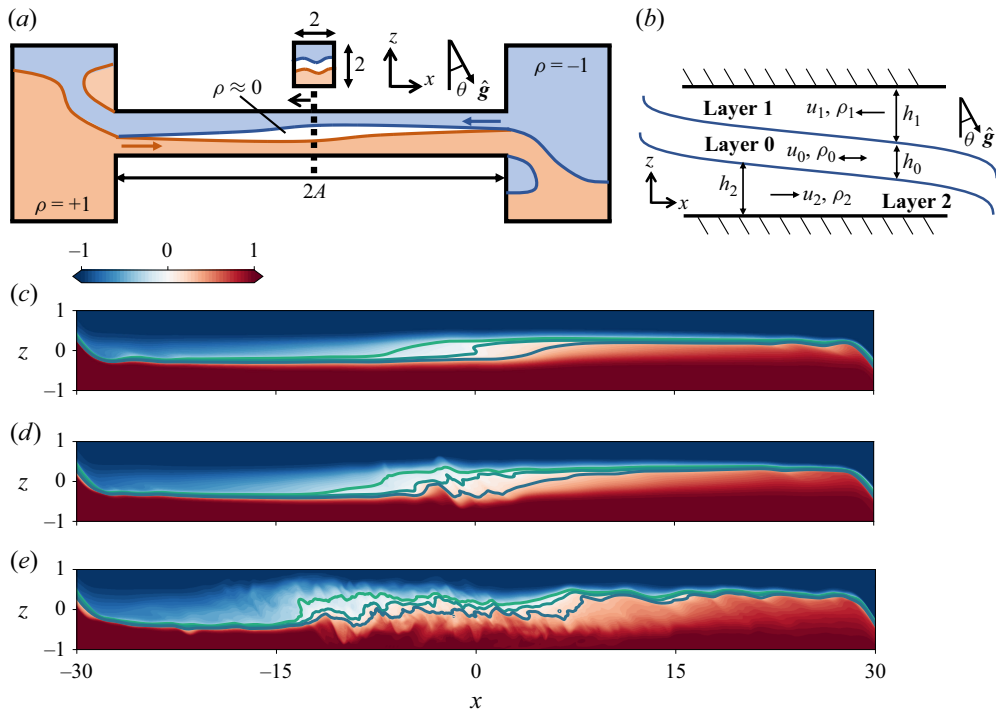


Figure 1. (a) Schematics of the SID flow configuration. The duct connecting two reservoirs (with different densities) has aspect ratios length-to-height $A = L_x^d/L_z^d = 30$ and width-to-height $B = L_y^d/L_z^d = 1$. (b) Schematics of the three-layer model. (c–e) Instantaneous snapshots of density fluctuation ρ around the reference value: (c) stationary wave (SW), (d) travelling wave (TW) and (e) intermittent turbulence (I).

has been much less discussed in the literature. Eckart (1961) and Ma (1981) discussed these long–short-wave interactions theoretically for idealised two-dimensional stratified flows and postulated them to occur when the group velocity of the short waves matches the phase speed of the long waves. However, evidence of such an interaction among long (and short) waves in a realistic situation requires prior knowledge of the variation of the flow on the scales comparable to the wavelength of the long waves, which is typically unavailable in these idealised cases.

The stratified inclined duct (SID) experiment (Meyer & Linden 2014) has been designed and employed to study stratified exchange flows in laboratory settings that resemble those in nature. In the SID experiment a long duct, which can be tilted at a small angle to the horizontal, connects two reservoirs filled with liquids with different densities (see figure 1). Flows in SID have been studied primarily in laboratory experiments (Lefauve & Linden 2020) and more recently by direct numerical simulations (DNS) (Zhu *et al.* 2023), where excellent agreement has been found.

The DNS of Zhu *et al.* (2023) were also analysed using two-layer hydraulics in Atoufi *et al.* (2023). The two-layer model is applicable when there is little mixing between the layers, and explains the onset of instability and turbulence. At high turbulence levels, however, the two-layer theory is unable to explain the features observed. Our first main objective in this paper is to analyse the three-layer internal hydraulics of exchange flows including both mixing and viscous effects. Our second main objective is to establish a link between the internal hydraulics and interfacial wave instabilities in three-layer flows by elucidating long–long and long–short-wave resonant interactions.

To achieve these aims, we start by describing our methodology in § 2, extending the SID flow DNS-data-driven approach of Atoufi *et al.* (2023) by diagnosing internal hydraulics in three-layer flows with improvements that include both mixing and viscous effects. The three-layer equations and hydraulic regimes are then discussed in § 3, including hydraulic controls as well as control mechanisms. Using this approach, we provide theoretical grounds to link three-layer hydraulics and instabilities in § 4. This allows us to quantify resonant interactions among long waves and between long and short waves, establishing a link with turbulence generation. Finally, we conclude in § 5.

2. Methodology: three-layer reduction of DNS data

2.1. DNS equations and data sets

The schematic of the SID flow is shown in figure 1(a). The configuration consists of a duct with internal height L_z^d , width L_y^d and length L_x^d connecting two reservoirs filled with fluids at densities $\rho_0 \pm \Delta\rho/2$. The entire configuration is inclined at an angle θ to the horizontal. The duct has a square cross-section and is 30 times longer than its height, with aspect ratios $A = L_x^d/L_z^d = 30$ and $B = L_y^d/L_z^d = 1$ for all cases in this study. Following Lefauve & Linden (2020), Zhu *et al.* (2023), we use duct half-height $L_z^d/2$, half-density difference $\Delta\rho/2$ for the non-dimensionalisation of lengths and density variations. The velocity is non-dimensionalised based on the averaged velocity difference of two oppositely propagating gravity currents $\Delta U/2 \equiv \sqrt{g' L_z^d}$, where $g' = g \Delta\rho/\rho_0$ is the reduced gravity. We note that this velocity difference across the density interface corresponds to the critical two-layer exchange flow with stability Froude number of unity $F_\Delta^2 = (\Delta U/2)^2/(g' L_z^d) = 1$ (Armi 1986; Dalziel 1991; Lawrence 1993; Atoufi *et al.* 2023). The non-dimensional parameters then become the tilt angle θ , Reynolds number ($Re = L_z^d \Delta U/4\nu$) and Prandtl number ($Pr = \nu/\kappa$), where ν and κ are momentum and mass diffusivity, respectively. The bulk Richardson number is, by definition,

$$Ri = \frac{g \frac{\Delta\rho}{2} \frac{L_z^d}{2}}{\rho_0 \left(\frac{\Delta U}{2}\right)^2} = \frac{1}{4}, \tag{2.1}$$

a constant in all cases in this study. The governing equations for the velocity field \mathbf{u} and density fluctuations ρ around the reference value are the forced Navier–Stokes equations under the Boussinesq approximation that takes the dimensionless form

$$\nabla \cdot \mathbf{u} = 0, \tag{2.2}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \frac{1}{Re} \nabla^2 \mathbf{u} + Ri \rho \hat{\mathbf{g}} - \mathbf{F}_u, \tag{2.3}$$

$$\frac{\partial \rho}{\partial t} + (\mathbf{u} \cdot \nabla) \rho = \frac{1}{Re Pr} \nabla^2 \rho - F_\rho, \tag{2.4}$$

where $\hat{\mathbf{g}} = (\sin \theta, 0, -\cos \theta)$ is the direction of gravity in the duct coordinate system (figure 1). The flow is driven in the streamwise $\hat{\mathbf{x}}$ direction and confined by solid walls in the spanwise direction $\hat{\mathbf{y}}$. The \mathbf{F}_u and F_ρ are forcing terms used to maintain the quasi-steady exchange flow, as described in Zhu *et al.* (2023). They are ONLY applied in a narrow region inside the reservoirs to sustain the flow without using excessively large reservoirs.

The governing equations are solved numerically using a compact sixth-order finite difference scheme for spatial derivatives and a third-order Adam-Bashforth scheme for time integration using the open-source solver Xcompact3D (Bartholomew *et al.* 2020)

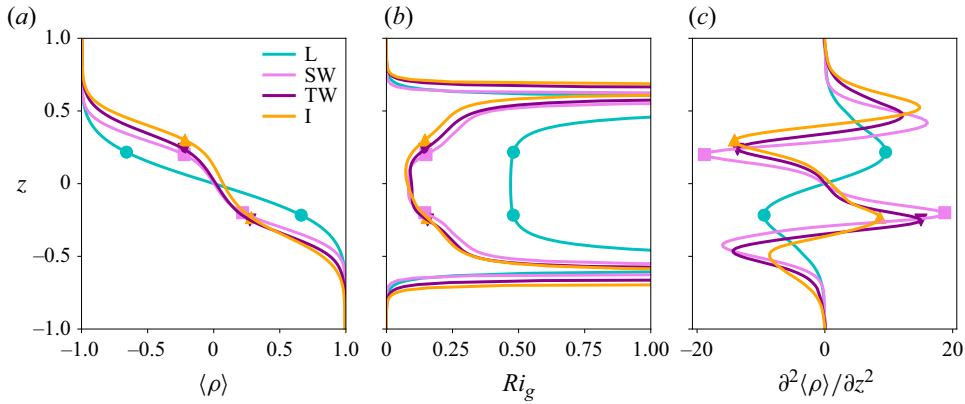


Figure 2. Layer splitting based on nearest turning points, illustrated by (a) mean density profiles, (b) gradient Richardson number Ri_g as defined in (2.8) and (c) the second derivative of the mean density profile, $\partial^2 \langle \rho \rangle / \partial z^2$. Markers indicate the nearest turning points to the mid-isopycnals where $\langle \rho \rangle = 0$ in each case.

modified to include F_ρ and F_u . For the full details of the numerical approach and comparison with experiments, the reader is referred to Zhu *et al.* (2023).

In this study we consider the laminar (L), stationary wave (SW), travelling wave (TW) and intermittent (I) datasets that correspond to $\theta = 2^\circ, 5^\circ, 6^\circ$ and 7° , respectively. All these cases are at the moderate Reynolds number $Re = 650$ and Prandtl number $Pr = 7$, representing thermal stratification in water. For detailed flow visualisations of the velocity and density fields corresponding to these regimes, readers are referred to Zhu *et al.* (2023), specifically their figure 5 and the supplementary materials, which include simulation movies.

2.2. Three-layer averaging

We reduce the dimensionality of the DNS data and cluster it into three distinct layers as sketched in figure 1(b). These distinct layers consist of a middle mixed layer between two other layers one above and another below as can be seen from the density snapshots in figure 1(c,d) for the TW and I cases. The three-layer structure of the density field for non-laminar cases is clear from the mean density profiles shown in figure 2(a). The slopes of the density profiles change in $-0.3 \leq z \leq 0.3$ for the SW, TW and I cases and, as a result of mixing, become closer to vertical than the L case in which the density profile can be reasonably considered as a two-layer profile.

We first average the velocity and density from simulation data over the y direction (denoted by $\langle \bullet \rangle_y = (1/2) \int_{-1}^1 \bullet dy$). We then locate the lower and upper interfaces from $\langle \rho \rangle_y$ that are denoted by $\eta_2(x, t)$ and $\eta_1(x, t)$, respectively (figure 1b).

To identify the bottom and top interfaces, we first locate the mid-isopycnal where $\langle \rho \rangle_y = 0$, as the dimensionless spanwise average density is $-1 \leq \langle \rho \rangle_y \leq 1$. The elevation of the mid-isopycnal is denoted by $\eta_0(x, t)$. The location of the bottom and top interfaces are marked using the inner turning points of $\langle \rho \rangle_y$, i.e. the points nearest to the mid-isopycnal where $\partial^2 \langle \rho \rangle_y / \partial z^2 = 0$ (figure 2). Those turning points with height $z = \eta_2(x, t) < \eta_0$ mark the location of the lower interface and those with height $z = \eta_1(x, t) > \eta_0$ mark the location of the upper interface.

We compute the layer-averaged velocities u_i , heights h_i and densities ρ_i (where $i = 0, 1, 2$ correspond to the middle, upper and lower layer, respectively) as height averages

$$h_1(x, t) = 1 - \eta_1(x, t), \quad u_1(x, t) = \langle \langle u \rangle_y \rangle_{z_1}, \quad \rho_1(x, t) = \langle \langle \rho \rangle_y \rangle_{z_1}, \quad (2.5)$$

$$h_0(x, t) = \eta_1(x, t) - \eta_2(x, t), \quad u_0(x, t) = \langle \langle u \rangle_y \rangle_{z_0}, \quad \rho_0(x, t) = \langle \langle \rho \rangle_y \rangle_{z_0}, \quad (2.6)$$

$$h_2(x, t) = 1 + \eta_2(x, t), \quad u_2(x, t) = \langle \langle u \rangle_y \rangle_{z_2}, \quad \rho_2(x, t) = \langle \langle \rho \rangle_y \rangle_{z_2}, \quad (2.7)$$

where the top layer average is $\langle \bullet \rangle_{z_1} = (1/h_1) \int_{\eta_1}^1 \bullet \, dz$, the middle layer average is given by $\langle \bullet \rangle_{z_0} = (1/h_0) \int_{\eta_2}^{\eta_1} \bullet \, dz$ and the bottom layer average is $\langle \bullet \rangle_{z_2} = (1/h_2) \int_{-1}^{\eta_2} \bullet \, dz$. We note that $z = 1$ and $z = -1$ are the non-dimensional height of the top and bottom walls, respectively.

As shown in figure 2(b), the gradient Richardson number defined based on velocity and density fields averaged over the duct length, cross-sections and time, i.e.

$$Ri_g \equiv -Ri \frac{\partial_z \langle \rho \rangle_{x,y,t}}{(\partial_z \langle u \rangle_{x,y,t})^2}, \quad (2.8)$$

is reasonably constant in the middle layer $\eta_1 < z < \eta_2$ and approximately equal to 0.5 for the laminar case and 0.15 for other cases. These values are consistent with the critical values for the local Richardson number observed in previous experimental studies (Lefauve & Linden 2020; Jiang *et al.* 2022, 2023).

2.3. Evolution of layer-averaged quantities

The velocity and heights of the three layers for the TW case are shown in figure 3 in $x-t$ (space-time) plots. We observe sudden changes in the layer thicknesses in the central region of the duct, indicated by the sharp transitions in contour colours, which we consider to be the locations of internal hydraulic jumps. The trajectories of these jumps are marked with two curves (dashed and dashed-dotted, corresponding to the upper and lower interfaces, respectively). We observe the region between these lines progressively extends towards both ends of the duct at later times. Consistent with the variations in layer heights, the layer velocities (u_0, u_1, u_2) also change abruptly at the same locations (see figure 3e-f). Note that the middle layer h_0 , which is initially located near the duct centre ($x \approx 0$) at $t = 80$ due to internal hydraulic jumps on the upper and lower interfaces, grows as the jumps propagate and eventually spread to occupy almost the entire duct by $t \approx 200$.

Figure 4 shows a schematic of the three-layer hydraulic jumps observed in the TW case. Near the left of the dashed line in figure 3(a,b), h_0 increases with a corresponding decrease in the upper layer thickness h_1 . The middle layer velocity u_0 is generally positive on the left side of the dashed line (associated with the location of the upper interface) but reduces to almost zero on the right side of the dashed line. Near the hydraulic jump, the velocities of the upper and middle layers have opposite signs (figure 3d,e) associated with a large difference in their respective heights (figure 3a,b). Such strong shear and asymmetry trigger instabilities and the upper interface becomes unstable. The bottom and middle layers have the same flow direction on the left side of the dashed line (figure 3d,f) and, therefore, have smaller shear; thus the lower interface is more stable here. A similar explanation can be provided for the instability of the lower interface by comparing the velocity and height differences between the middle and lower layers at the neighbourhood of the right hydraulic jump (dashed-dotted lines) as illustrated in figure 4. These jumps in the layer heights and velocities and the associated turbulence generation are further linked to three-layer internal hydraulics in the next section.

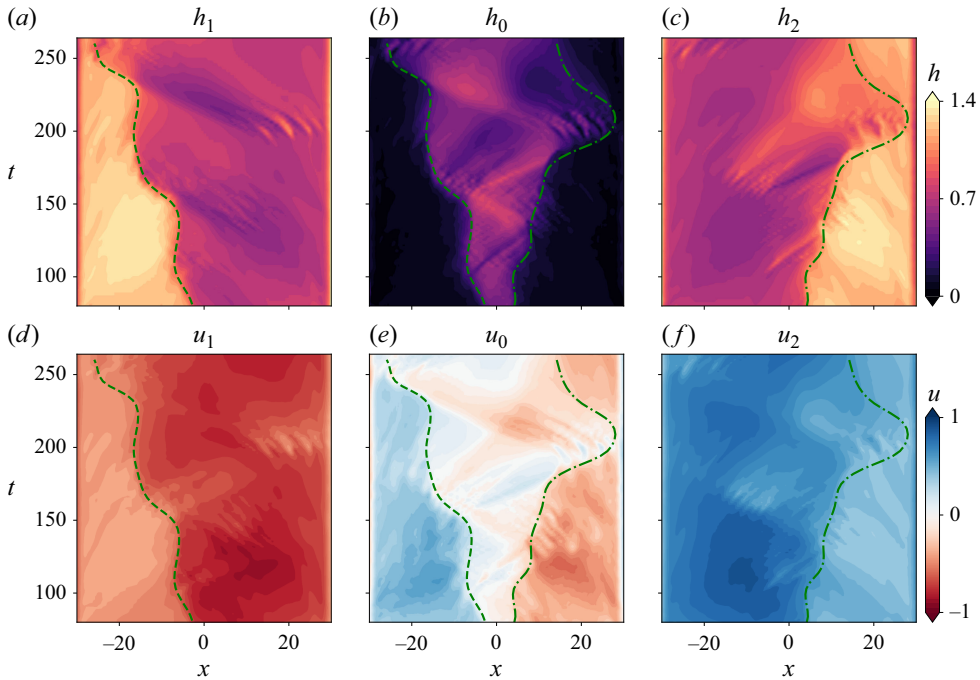


Figure 3. Space–time (x – t) plots of layer heights and velocities obtained from DNS for the TW case, assuming a three-layer model: (a, b, c) represent the heights and (d, e, f) the velocities of the upper, middle and lower layers, respectively. Dashed and dash-dotted lines indicate abrupt changes in the heights of the upper and lower edges of the middle layer, respectively, corresponding to identical changes in velocities across all panels.

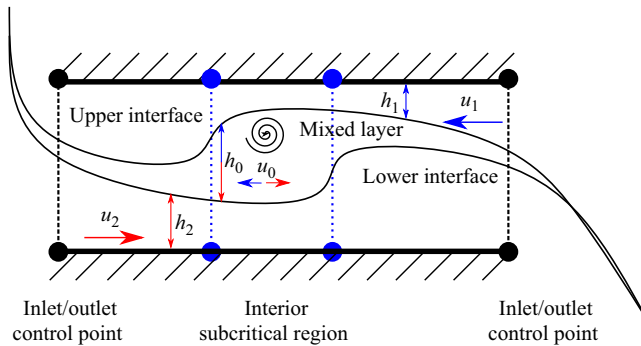


Figure 4. A schematic of the internal hydraulic jumps in SID flows.

3. Three-layer hydraulics: characteristics, regimes and control mechanisms

In the previous section we discussed temporal and along-duct variations of layer heights and velocities (mainly for the SW and TW cases) and postulated a possible link between such variation in layer properties and instability and the transition to turbulence. In this section we introduce viscous, non-hydrostatic three-layer equations that describe the internal hydraulics of stratified turbulent exchange flows in § 3.1. These equations also govern the dynamics of DNS data using the three-layer averaging procedure described in § 2.2. Building upon the characteristics of the three-layer equations, we demonstrate information propagation in § 3.2 and identify three-layer hydraulic control and hydraulic

transitions in § 3.3. We then show their strong correlation to turbulence generation in § 3.3.3. In § 3.4 we discuss the regularisation of the three-layer exchange flow near the hydraulic control points (i.e. hydraulic control mechanisms).

3.1. Three-layer evolution equations

To identify mechanisms governing the evolution of three-layer exchange flow, we perform a layer-wise integration of (2.2)–(2.4) as described in § 2.2, noting that $F_u = 0$, $F_p = 0$ inside the duct where DNS data were collected. This procedure leads to the viscous three-layer equations and can be written in the general vector form as

$$\mathbf{C} \frac{\partial \mathbf{q}}{\partial t} + \mathbf{A} \frac{\partial \mathbf{q}}{\partial x} = \mathbf{B} \frac{\partial s}{\partial x} + \boldsymbol{\tau} \mathbf{q} + \mathcal{F}, \tag{3.1}$$

where the state vector is defined as $\mathbf{q} = [u_1, u_0, u_2, h_1, h_0, h_2]^T$, recalling that the subscripts identify the velocities and heights of the upper (1), middle (0) and lower (2) layer, respectively, as sketched in figure 1(b). The coefficient matrices on the left-hand side of (3.1) are

$$\mathbf{A} = \begin{pmatrix} -u_1 & u_0 & 0 & 0 & g_1 & g_1 \\ 0 & -u_0 & u_2 & 0 & 0 & g_2 \\ -h_1 & h_0 & 0 & -u_1 & u_0 & 0 \\ 0 & -h_0 & h_2 & 0 & -u_0 & u_2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ h_1 & h_0 & h_2 & u_1 & u_0 & u_2 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.2}$$

where $g_1 = Ri \cos \theta (\rho_0 - \rho_1)$ and $g_2 = Ri \cos \theta (\rho_2 - \rho_0)$ are the dimensionless reduced gravity at the upper and lower interfaces, respectively. The forcing terms, $\mathbf{B} \partial s / \partial x$ and $\boldsymbol{\tau} \mathbf{q}$, are due to buoyancy and viscous stresses and are discussed in detail in § 3.4. The gravity tensor reads as

$$\mathbf{B} = \begin{pmatrix} g_1 \tan \theta & -g_1 & \mathbf{0}_{1,4} \\ g_2 \tan \theta & -g_2 & \mathbf{0}_{1,4} \\ \mathbf{0}_{4,1} & \mathbf{0}_{4,1} & \mathbf{0}_{4,4} \end{pmatrix}, \tag{3.3}$$

where $\mathbf{0}_{m,n}$ denotes the block null matrix of the size $m \times n$ and $s = [x, b, 0, 0, 0, 0]^T$ is the vector associated with geometry change. Here $b(x)$ denotes the elevation of the bottom wall where $b(x) = -1$ throughout the duct but abruptly changes to $b(x) = -2$ at the inlet/outlet. The model for the viscous stress tensor for three-layer hydraulics is a sparse 6×6 matrix,

$$\boldsymbol{\tau} = \frac{2}{Re} \begin{pmatrix} \frac{1}{h_0 h_1} - \frac{f_a}{h_1} & \frac{h_1 - h_0}{h_0 h_1 (h_1 + h_0)} - \frac{1}{h_0 (h_0 + h_2)} & \frac{1}{h_0 (h_0 + h_2)} & \mathbf{0}_{1,3} \\ \frac{1}{h_0 (h_1 + h_0)} & \frac{h_1 + h_0 + h_2}{h_0 (h_1 + h_0) h_2} & \frac{h_2 - h_0}{h_0 h_2 (h_0 + h_2)} - \frac{f_b}{h_2} & \mathbf{0}_{1,3} \\ \mathbf{0}_{4,1} & \mathbf{0}_{4,1} & \mathbf{0}_{4,1} & \mathbf{0}_{4,3} \end{pmatrix}, \tag{3.4}$$

where f_a and f_b are the slip-with-friction coefficients at the upper and lower boundaries. These coefficients are computed based on $f \langle u \rangle_y = \partial \langle u \rangle_y / \partial z$ at the bottom and top boundaries from DNS data. For a free-slip condition $f = 0$ and for the no-slip condition,

f is proportional to Re . The forcing term $\mathcal{F} = [\mathcal{F}_1, \mathcal{F}_0, \mathcal{F}_2, 0, 0, 0]^T$ is due to the non-hydrostatic pressure gradient and is determined from DNS data (note that this contribution is weaker than the hydrostatic pressure gradient in all cases considered here; see Atoufi *et al.* 2023). If $\mathcal{F} = 0$, (3.1) reduce to the viscous shallow water equations.

3.2. Information propagation

We now investigate, using (3.1), the directions in which information is carried by long interfacial waves in a three-layer configuration, dictating the hydraulic regime. Consider a left eigenvector \mathbf{v} and eigenvalue λ associated with the coefficient matrix pair (\mathbf{A}, \mathbf{C}) such that

$$\mathbf{v}^H \mathbf{A} = \lambda \mathbf{v}^H \mathbf{C}, \tag{3.5}$$

where the superscript H denotes the Hermitian transpose. Multiplying (3.1) by \mathbf{v}^H and rewriting $\boldsymbol{\tau} = \mathbf{C}\mathbf{D}$ yields

$$\mathbf{v}^H \mathbf{C} \underbrace{\left(\frac{\partial \mathbf{q}}{\partial t} + \lambda(x, t) \frac{\partial \mathbf{q}}{\partial x} - \mathbf{D}\mathbf{q} \right)}_{\text{Information propagation with loss/gain}} = \mathbf{v}^H \underbrace{\left(\mathbf{B} \frac{\partial s}{\partial x} + \mathcal{F} \right)}_{\text{Modulations}}, \tag{3.6}$$

where

$$\mathbf{D} = \begin{pmatrix} \mathbf{d} & \mathbf{0}_{3,2} & \mathbf{0}_{3,2} \\ \mathbf{0}_{3,2} & \mathbf{d} & \mathbf{0}_{3,2} \end{pmatrix} \boldsymbol{\tau}, \quad \text{such that } \mathbf{d} = \frac{1}{3} \begin{pmatrix} -2 & -1 \\ 1 & -1 \\ 1 & 2 \end{pmatrix}, \tag{3.7}$$

is the diffusion matrix obtained by a singular value decomposition of $\mathbf{C} = \mathbf{U}\mathbf{Q}\mathbf{V}^H$, where $\mathbf{U}\mathbf{U}^H = \mathbf{V}^H\mathbf{V} = \mathbf{I}$ and \mathbf{Q} is a diagonal matrix leading to $\mathbf{D} = \mathbf{V}\mathbf{Q}^{-1}\mathbf{U}^H\boldsymbol{\tau}$.

Recall that $\mathbf{D}\mathbf{q}$ is solely due to viscous effects and when absorbed in the left-hand side of (3.6) it can be viewed as a mechanism affecting the propagation of information contained in \mathbf{q} .

Equation (3.6) is particularly useful in explaining various mechanisms affecting the propagation of information in stratified turbulent exchange flows. In the limit where viscous stresses are neglected ($\mathbf{D} = \mathbf{0}$), non-hydrostatic pressure gradients are ignored ($\mathcal{F} = \mathbf{0}$), in the interior region when the configuration is horizontal ($\mathbf{B} \partial s / \partial x = \mathbf{0}$) then (3.6) will be homogeneous.

In the homogeneous limit, the coordinate transformation $2\xi = t + x/\lambda$ ($\lambda \neq 0$) yields

$$\mathbf{v}^H \mathbf{C} \left(\frac{\partial \mathbf{q}}{\partial t} + \lambda(x, t) \frac{\partial \mathbf{q}}{\partial x} \right) = \mathbf{v}^H \mathbf{C} \frac{d\mathbf{q}}{d\xi} \left(1 - \frac{\overbrace{\partial_t \lambda + \lambda \partial_x \lambda}^{\approx 0}}{2\lambda^2} \right) = \mathbf{0}, \tag{3.8}$$

and a combination of flow variables $\mathbf{v}^H \mathbf{C} d\mathbf{q} / d\xi = \mathbf{0}$ that is conserved to leading order. Note that here the space–time variation of λ is assumed to be small compared with that of \mathbf{q} . Therefore, interfacial waves travel ‘freely’ along the characteristic curves defined by ξ with characteristic velocities λ (simply referred to as ‘characteristics’).

In the inhomogeneous case, (3.6) suggests that information is also carried by similar characteristics λ but with two modifications. Once the viscous stresses are considered the information propagation along the characteristic curves is accompanied by the loss due to a viscous damping effect (momentum loss) or gain due to mixing (thickening of the middle layer) both originating from the $\mathbf{D}\mathbf{q}$ term. The gravitational forcing due

to tilting and geometrical variation in $\mathbf{B} \partial s / \partial x$ as well as the non-hydrostatic pressure forcing in \mathcal{F} do not affect the direction of propagation (i.e. modulations) as they do not involve the information content \mathbf{q} . Note that in the inhomogeneous viscous case, there is no conservation of any combination of flow variables along the characteristic curves ξ . Nevertheless, we employ the term ‘characteristic’, drawing its interpretation from the homogeneous inviscid case.

The preceding analysis assumed a real characteristic velocity λ . More generally, λ may be complex, in which case the interpretation of characteristics must be revisited. As shown in Appendix D, the real part λ^R continues to determine the direction of information propagation to leading order, extending the interpretation of real-valued characteristics.

3.2.1. Analytical solution for the characteristics

The characteristics λ of the three-layer model are obtained by solving for the eigenvalues in (3.5) that requires $\det(\mathbf{A} - \lambda \mathbf{C}) = 0$ for a non-trivial solution leading to the characteristic polynomial

$$\sum_{n=0}^4 a_n \lambda^n = 0, \tag{3.9}$$

with the coefficients

$$\begin{aligned} a_4 &= (h_0 + \widehat{h}), \\ a_3 &= -2 [\widehat{h} u_0 + (u_1 + u_2)h_0 + \widehat{h} \widehat{u}], \\ a_2 &= \widehat{h} u_0^2 + 4\widehat{h} \widehat{u} u_0 + g_1 h_1 (h_0 + h_2)(F_1^2 - 1) + g_2 h_2 (h_0 + h_1)(F_2^2 - 1) \\ &\quad + 4 h_0 u_1 u_2, \\ a_1 &= -2\widehat{h} \widehat{u} u_0^2 - 2g_1 h_1 h_2 (F_1^2 - 1) u_0 - 2g_2 h_1 h_2 (F_2^2 - 1) u_0 - 2 h_0 u_1 u_2 (u_1 + u_2) \\ &\quad + 2 h_0 \widehat{h} \widehat{u} \widehat{g}, \\ a_0 &= \det \mathbf{A}. \end{aligned} \tag{3.10}$$

In these coefficients we introduced the height $\widehat{h} = h_1 + h_2$, velocity $\widehat{u} = (h_1 u_2 + h_2 u_1) / \widehat{h}$ and reduced gravity $\widehat{g} = (g_1 h_1 u_2 + g_2 h_2 u_1) / (\widehat{h} \widehat{u})$. We also used the Froude numbers of the bottom, middle and top layers, respectively,

$$F_2 = \frac{u_2}{\sqrt{g_2 h_2}}, \quad F_0 = \frac{u_0}{\sqrt{\frac{g_1 g_2}{g_1 + g_2} h_0}}, \quad F_1 = \frac{u_1}{\sqrt{g_1 h_1}}. \tag{3.11}$$

The characteristics have the form

$$\begin{aligned} \lambda_1 &= \bar{\lambda} - \delta_1 \lambda - \delta_2 \lambda, & \lambda_2 &= \bar{\lambda} - \delta_1 \lambda + \delta_2 \lambda, \\ \lambda_3 &= \bar{\lambda} + \delta_1 \lambda - \delta_3 \lambda, & \lambda_4 &= \bar{\lambda} + \delta_1 \lambda + \delta_3 \lambda, \end{aligned} \tag{3.12}$$

where $\bar{\lambda} = -a_3 / 4a_4$ is the convective velocity and

$$\begin{aligned} \delta_1 \lambda &= \frac{1}{2} \sqrt{-\frac{2}{3} m_1 + \frac{1}{3a_4} \left(m_3 + \frac{\Delta_0}{m_3} \right)}, \\ \delta_2 \lambda &= \frac{1}{2} \sqrt{-4(\delta_1 \lambda)^2 - 2m_1 + \frac{m_2}{\delta_1 \lambda}}, & \delta_3 \lambda &= \frac{1}{2} \sqrt{-4(\delta_1 \lambda)^2 - 2m_1 - \frac{m_2}{\delta_1 \lambda}} \end{aligned} \tag{3.13}$$

are the phase speeds, where

$$m_1 = \frac{8a_4a_2 - 3a_3^2}{8a_4^2}, \quad m_2 = \frac{a_3^2 - 4a_4a_3a_2 + 8a_4^2a_1}{8a_4^3}, \quad m_3 = \left(\frac{\Delta_1^2 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2} \right)^{\frac{1}{3}}, \tag{3.14}$$

$$\Delta_0 = a_2^2 - 3a_3a_1 + 12a_4a_0, \quad \Delta_1 = 2a_2^3 - 9a_3a_2a_1 + 27a_3^2a_0 + 27a_4a_1^2 - 72a_4a_2a_0.$$

Since $(\delta_1\lambda^R, \delta_2\lambda^R, \delta_3\lambda^R) \geq 0$, we realise that $\lambda_1^R \leq \lambda_2^R$, $\lambda_3^R \leq \lambda_4^R$ and $\lambda_1^R \leq \lambda_4^R$, where the superscript R denotes the real component. Consequently, in a reference frame moving at $\bar{\lambda}$, an observer would see two oppositely propagating interfacial waves, which are associated with the eigenvalues $\lambda_{1,2}$ and $\lambda_{3,4}$.

To justify this association, we note that the terms $\delta\lambda = -\delta_1\lambda \mp \delta_2\lambda$ and $\delta\lambda = \delta_1\lambda \mp \delta_3\lambda$ represent the effective phase speeds of interfacial waves in a three-layer flow. The structure of these expressions suggests that $\lambda_{1,2}$ and $\lambda_{3,4}$ correspond to long waves propagating along the upper and lower interfaces, respectively. Moreover, since the coefficients in (3.10) are real, if any characteristics become imaginary, they must appear in complex conjugate pairs. Complex conjugate eigenvalues correspond to physically linked interfacial waves, reinforcing the idea that each pair (λ_1, λ_2) and (λ_3, λ_4) can be naturally associated with a single interface. Thus, a key observation from (3.12) is that the eigenvalues $\lambda_{1,2}$ and $\lambda_{3,4}$ can be identified with long waves at the upper and lower interfaces, respectively. We note this labelling does not imply that the interfaces are completely decoupled, as their characteristics inherently depend on the velocities and heights of all three layers.

3.2.2. A special case: pure exchange flow with stagnant middle layer

Of particular interest is identifying the conditions under which interfacial waves become unstable. To make progress on this question, we examine a special asymmetric case where $u_2 = -u_1$, $h_1 = h_2$ and $g_1 = g_2$. The zero (barotropic) net flow condition implies that $u_0 = -(u_1h_1 + u_2h_2)/h_0 = 0$ and the coefficient a_3 vanishes, leading to zero convective velocity ($\bar{\lambda} = 0$). This corresponds to the simplest three-layer exchange flow, where the upper and lower layers are symmetric in thickness and stratification, while the middle layer remains stationary. This case is also highly relevant to the SID flow, particularly near the centre of the duct. Under these conditions, the expressions for characteristics in (3.12) simplify significantly (since $F_1 = F_2$), yielding

$$\begin{aligned} \lambda_{1,2} &= \mp \frac{\sqrt{g_1h_1}}{\sqrt{h_0 + 2h_1}} \sqrt{h_0(1 + F_1^2) + h_1(1 - F_1^2) - \sigma}, \\ \lambda_{3,4} &= \mp \frac{\sqrt{g_1h_1}}{\sqrt{h_0 + 2h_1}} \sqrt{h_0(1 + F_1^2) + h_1(1 - F_1^2) + \sigma}, \end{aligned} \tag{3.15}$$

where $\sigma = \sqrt{4h_0h_1F_1^2(1 - F_1^2) + 4h_0^2F_1^2 + h_1^2(1 - F_1^2)^2}$. Clearly, in this case, if we further restrict the system to $F_1^2 = 1$ then $\lambda_{1,2} = 0$ and $\lambda_{3,4} = \mp 2\sqrt{g_1h_0h_1}/\sqrt{h_0 + 2h_1}$. Since waves in the upper and lower layers become decoupled in this limit, this result supports our approach of associating each eigenvalue pair $\lambda_{1,2}$ and $\lambda_{3,4}$ with a specific interface. However, we note that interfacial waves are coupled where $\bar{\lambda} \neq 0$ in general.

Equation (3.15) reveals that the onset of complex characteristics, indicative of long-wave instability, can be anticipated by identifying the condition under which σ vanishes. Setting $\sigma = 0$ yields the critical Froude number thresholds

$$F_{\Delta\mp}^2 = 1 + \frac{2r(r - 1) \mp 2r^{3/2}\sqrt{r + 2}}{4r - 1}, \tag{3.16}$$

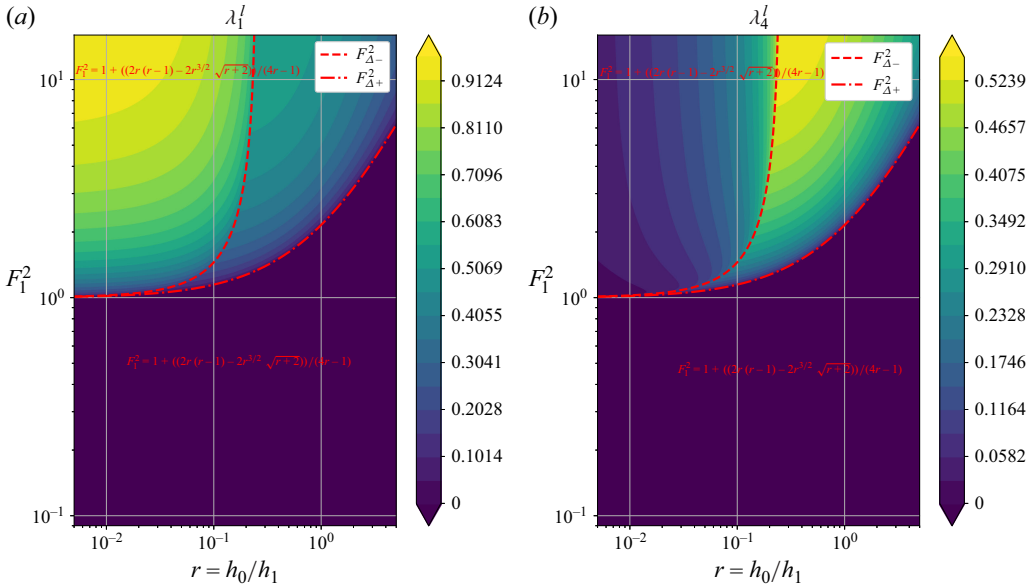


Figure 5. The onset of long-wave instability in the pure exchange flow case visible where (a) $\lambda_1^I > 0$ and (b) $\lambda_4^I > 0$. Here, F denotes the Froude number of the upper layer and r represents the ratio of middle-to-upper/lower layer heights.

where $r = h_0/h_1$ denotes the ratio of the middle layer depth to the upper (or lower) layer depth. Figure 5 shows λ_1^I and λ_4^I as functions of F_1^2 for the pure exchange flow. The critical Froude numbers predicted by (3.16) are shown as red dashed lines. We observe that, for $F_1^2 \geq F_{\Delta+}^2$, the characteristics become complex, and the largest values of λ_4^I occur within the range $F_{\Delta+}^2 \leq F_1^2 \leq F_{\Delta-}^2$. Moreover, the magnitudes of λ_1^I and λ_4^I change abruptly across the $F_{\Delta-}^2$ curve. We identify $F_{\Delta-}^2$ as the bifurcation Froude number and $F_{\Delta+}^2$ as the marginal-stability Froude number for the pure exchange flow case.

The velocity difference between the upper and middle layers is given by $u_1^2 = F_1^2 g_1 h_1$, since $u_0 = 0$ in the pure exchange flow. Hence, for $F_1^2 \geq F_{\Delta+}^2$, the velocity shear between the upper and middle layers, and between the middle and lower layers, becomes sufficiently strong to trigger long-wave instability on each interface separately. As r increases, the middle layer thickens relative to the upper layer, causing the upper layer to thin and accelerate. This leads to larger velocity differences across the interfaces, particularly between the faster upper layer and the stagnant middle layer, leading to instability once $F_1^2 = F_{\Delta+}^2$. This behaviour is analogous to that observed in two-layer flows, where sufficiently strong interfacial shear also leads to long-wave instability.

The condition $\sigma = 0$ corresponds to $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$, which implies that the interfacial waves on the lower and upper interfaces have equal phase speeds and can resonate. Instability may therefore arise from this resonance. In the more general case with non-zero convective velocity, however, the identification of instability is less straightforward, a point revisited below and addressed in detail in § 4.

3.2.3. Criticality of interfaces

In a two-layer configuration, the sign of the real component of the characteristics sets the direction of information propagation and the hydraulic regime of the flow, i.e.

subcritical, critical or supercritical regime (Long 1956; Dalziel 1991; Atoufi *et al.* 2023). The flow is subcritical when $\bar{\lambda} < \delta\lambda^R$ and information propagates in both leftward (towards decreasing x) and rightward (towards increasing x) directions, supercritical when $\bar{\lambda} > \delta\lambda^R$ and information propagates only either rightward or leftward and critical when $\bar{\lambda} = \delta\lambda^R$.

In a three-layer configuration, complications arise as the upper and lower layers can individually exhibit different hydraulic regimes. As we found above, $\lambda_{1,2}$ and $\lambda_{3,4}$ are respectively associated with the upper and lower interface. Therefore, the necessary condition for the upper (respectively lower) layer to be supercritical is $\lambda_1^R \lambda_2^R > 0$ (respectively $\lambda_3^R \lambda_4^R > 0$). Information on each interface thus propagates in a particular direction, depending on the signs of the characteristics (Sannino *et al.* 2007, 2009).

The characteristic pairs $\lambda_1^R \lambda_2^R$ and $\lambda_3^R \lambda_4^R$ determine criticality with respect to the individual layers, not the full three-layer system. For example, the flow may be supercritical in both layers if $\lambda_1^R \lambda_2^R > 0$ and $\lambda_3^R \lambda_4^R > 0$, yet remain subcritical overall, e.g. if $\lambda_1^R, \lambda_2^R < 0$ and $\lambda_3^R, \lambda_4^R > 0$. The three-layer flow is fully supercritical only when all layers are hydraulically supercritical and the corresponding characteristics have the same sign, so that information propagates in a single direction. It is fully subcritical when all layers are subcritical, or when the layers are supercritical but the characteristic signs differ, indicating opposing directions of information propagation. A critical condition arises when at least one characteristic pair satisfies $\lambda_i^R \lambda_j^R = 0$, at which point information propagation is blocked and internal hydraulic control is established.

3.2.4. Critical middle layer thickness

The emergence of imaginary components, as conjugate pairs, in the characteristics implies the onset of instability of long waves as established in Appendix A. Consequently, we can predict critical values of the layer velocities and heights that cause interfaces to be unstable to disturbances of the given wavenumber k by examining the coefficients b_i of the quartic stability equation (A10) given below,

$$\begin{aligned} b_4 &= \sinh(k(h_0 + \hat{h})), \\ b_3 &= \sinh(k(h_0 - \tilde{h}))(u_0 - u_1) - b_4(u_1 + u_2 + 2u_0) + \sinh(k(h_0 + \tilde{h}))(u_0 - u_2), \\ b_2 &= u_0^2(\alpha_1 + \alpha_2 + 6\alpha_4) + u_1^2(\alpha_1 + \alpha_3) + u_2^2(\alpha_2 + \alpha_3) + 4u_0(u_1\alpha_1 + u_2\alpha_2) + 4u_1u_2\alpha_3 \\ &\quad + \gamma_1 + \gamma_2, \\ b_1 &= (1 - 2\sigma_1^2)(\gamma_7 + \gamma_9) - \sigma_1\sigma_2\gamma_8, \\ b_0 &= u_0^4\gamma_3 + 2u_0^4\sigma_1\beta_1\sigma_2 + 2u_1^2u_2^2\beta_4\sigma_1\sigma_2 + u_2^2\gamma_4 + u_1^2\gamma_5 + \gamma_6, \end{aligned} \tag{3.17}$$

and to long waves by examining a_i in (3.10) where $\tilde{h} = h_1 - h_2$ and α_i, σ_i and γ_i are defined in Appendix A.1. For example, the necessary condition guaranteeing only real roots is that b_i should be all positive or all negative (based on the Routh–Hurwitz stability criterion (Shinners 1998)). As wavenumbers $k \geq 0$, then $b_4 \geq 0$ and this condition leads to one of the $b_3, b_2, b_1, b_0 < 0$ for all k for instability to happen.

In fact, the number of sign changes in b_i yields the number of complex roots and potential instability. As λ (i.e. the long-wave limit of the phase speed c of disturbance waves) appears in the form of complex conjugates, there are four complex roots and $a_{3,2,1,0}$ change sign (note a_4 is constant inside the duct). In the limit of long waves, using the $b_3 = a_3 < 0$ condition and incorporating the zero net flow condition $Q = u_0h_0 + u_1h_1 + u_2h_2 = 0$ to eliminate u_0 in a_3 yields a restriction for the height of the middle layer that $-2(h_0 + \tilde{h})(h_0(u_1 + u_2) - h_1u_1 - h_2u_2)/h_0 < 0$ for instability to occur.

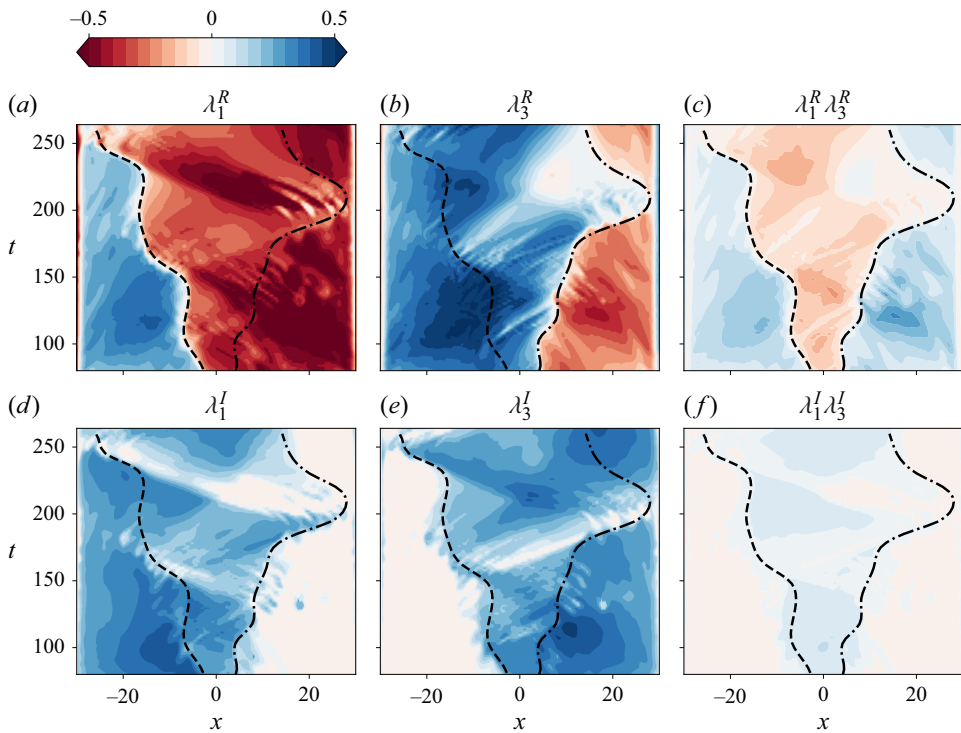


Figure 6. Selected eigenvalues of the three-layer model for the TW case ($Re = 650, \theta = 6^\circ$): (a,d) real and imaginary components of the characteristic λ_1 for the upper interface, (b,e) λ_3 for the lower interface and (c,f) product of the real and imaginary components of λ_1 and λ_3 . The dashed and dash-dotted lines correspond to those in figure 3.

This leads to

$$h_0^{cr} = \frac{u_1 h_1 + u_2 h_2}{u_1 + u_2} \tag{3.18}$$

as the critical height for the middle layer. Therefore, when $u_2 < |u_1|$ (i.e. an interface with a positive slope) and $h_0 > h_0^{cr}$, the three-layer exchange flow is become unstable to long waves and may be supercritical. In the limit of short waves, where $k \gg 1$, the sign of b_3 depends on the coefficients of the largest k (positive) when expanding b_3 with respect to k . Specifically, when $u_2 h_0 + u_1 h_1 + u_2 h_2 > 0$, we have $b_3 < 0$, satisfying the necessary condition. This leads to the condition $h_0 > (|u_1| h_1 - u_2 h_2) / u_2$, which provides a smaller lower bound for h_0 and a smaller critical middle layer thickness than in (3.18). Therefore, it is expected that short waves will become unstable first. Nonetheless, it is the long waves (hydraulic effects) that set the interface slopes, and thus, the upper and lower layer flow rates $u_1 h_1$ and $u_2 h_2$, which then determine the value of h_0 that results in $b_3 < 0$ and the onset of the short-wave instabilities.

3.2.5. Application to the data

The three-layer model, yielding layer characteristics, is applied to the layer-averaged DNS data in figure 6. The space–time evolution of λ_1 and λ_3 in the TW case are shown with $x-t$ plots. The dashed lines are reproduced from the layer-averaged quantities in figure 3, indicating the drastic changes in the layer statistics. Noticeably, the left and right lines are, respectively, in alignment with λ_1^R and $\lambda_3^R \approx 0$, which indicates the emergence of critical

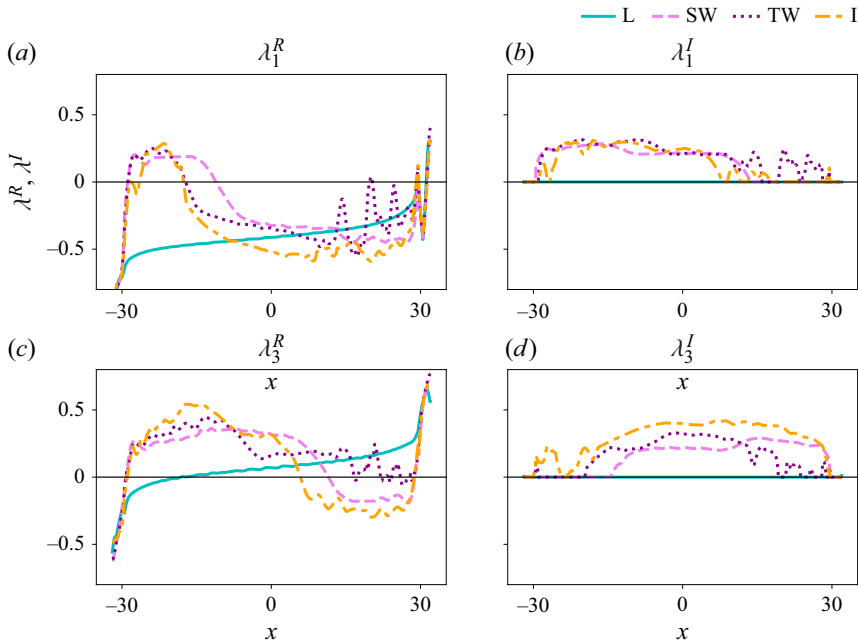


Figure 7. Instantaneous characteristics of the interfacial waves on the upper and lower interfaces in DNS, diagnosed using the three-layer hydraulic analysis. The real components of the characteristics, representing the phase speed of the upper and lower interfacial waves (λ_1^R and λ_3^R), are shown in (a,c). The imaginary components, representing the growth rate of the upper and lower interfacial waves (λ_1^I and λ_3^I), are shown in (b,d).

points in the upper and lower layer interfaces. The flow becomes critical on these lines, which is also evidenced by the appearance of hydraulic jumps as illustrated by the sudden changes of the corresponding layer heights in figure 3. In addition, the three-layer model predicts one control point near the inlet/outlet of the duct due to geometry changes, which agrees with the two-layer model in Atoufi *et al.* (2023).

In figure 6(c), outside of the critical lines, the real components of the characteristics have the same sign and are positive/negative on the left/right side of the duct, respectively. Therefore, the information of the three-layer system can only propagate from the critical point near the inlet/outlet toward the duct centre. In between the critical lines, $\lambda_1^R \lambda_3^R < 0$ and the flow is subcritical. Waves in this region can propagate in either direction. This region is where the thickening of the middle layer occurs (see figure 3).

The imaginary components of characteristics correspond to the growth rate of the long interfacial waves (as discussed in Appendix A). As indicated by figure 6(d,e), the long waves at the upper interface become unstable ($\lambda_1^I > 0$) in the left and middle regions of the duct, whereas the long waves at the lower interface are unstable in the middle and right regions ($\lambda_3^I > 0$) where the superscript I refers to the imaginary component of λ . While the stability of individual interfaces can be deduced from the associated characteristics, the stability of the three-layer flow system is less straightforward and requires consideration of resonance between interfaces, which will be discussed in § 4.

The real and imaginary components of the instantaneous characteristics are shown in figure 7 at $t = 120$ for the L, SW, TW and I cases, to complement the space–time distribution of the characteristics, which was previously given only for the TW case. The instantaneous λ^R and λ^I of all non-laminar cases follow a similar trend. The imaginary

component in figure 7(b,d) shows that in the L case, $\lambda^I \approx 0$ for both the upper and lower interfaces, whereas in the SW, TW and I cases, $\lambda_1^I, \lambda_3^I > 0$ in most places inside the duct.

The application of the three-layer theory to the DNS data suggests that three-layer characteristics are not only useful in quantifying information propagation but also in informing us about regions where turbulence and mixing occur. We detail this connection between characteristics and turbulence next.

3.3. Hydraulic regime transitions and turbulence

Here we identify an equivalent Froude number of a three-layer flow that predicts hydraulic regime transition and control in a three-layer exchange flow in § 3.3.2. Then, we use this three-layer Froude number to show how hydraulic control leads to turbulence generation in § 3.3.3.

3.3.1. Critical exchange flow regime

To define hydraulic control, we use the steady-state solution of (3.1). For $\mathcal{F} = 0$, the steady solution of (3.1) can be found by multiplying the inverse matrix $\mathbf{A}^{-1} = \text{adj}(\mathbf{A})/\det \mathbf{A}$ on both sides leading to

$$\frac{\partial \mathbf{q}}{\partial x} = \frac{1}{\det \mathbf{A}} \text{adj}(\mathbf{A}) \mathbf{B} \frac{\partial s}{\partial x} + \frac{1}{\det \mathbf{A}} \text{adj}(\mathbf{A}) \tau \mathbf{q}, \quad (3.19)$$

where $\text{adj}(\mathbf{A})$ represent the transpose of the co-factor matrix of \mathbf{A} . A non-trivial steady solution $\mathbf{q}(x)$ of (3.1) requires

$$\det \mathbf{A} = g_1 g_2 h_1 h_0 h_2 (G - 1) \neq 0, \quad (3.20)$$

where G is the composite Froude number defined based on the individual layers

$$G = 1 + (\epsilon_1 F_1^2 + \epsilon_2 F_2^2 - 1) F_0^2 + (F_1^2 - 1) (F_2^2 - 1), \quad (3.21)$$

where $\epsilon_1 = g_1/(g_1 + g_2)$, $\epsilon_2 = g_2/(g_1 + g_2)$ are the ratio of cross-interface to total density differences.

The three-layer flow is hydraulically controlled when $h_0 = 0$ or $G = 1$ (i.e. $\det \mathbf{A} = 0$). The condition $h_0 = 0$ is trivial as it removes the third layer. Generally, the hydraulic control condition incorporating (3.21) defines a critical state in the three-layer exchange flow, where

$$F_0^2 = \frac{(F_1^2 - 1) (F_2^2 - 1)}{(1 - \epsilon_1 F_1^2 - \epsilon_2 F_2^2)}, \quad (3.22)$$

yielding $G = 1$. This equation imposes a fundamental constraint on the middle layer, determining its velocity and thickness based on the Froude numbers of the upper and lower layers, thereby fully characterising its dynamics under hydraulic control. Determining whether a three-layer exchange flow beyond the critical state is subcritical or supercritical requires linking (3.21) to the characteristics of long interfacial waves, as discussed in the next section. Nonetheless, (3.22) allows the exploration of the possible combinations of supercritical and subcritical flows in each layer even though the overall three-layer system remains in a critical state. Figure 8(a) shows the middle layer Froude number F_0^2 at the critical state, computed from (3.21) for $\epsilon_1 = \epsilon_2 = 0.5$, as a function of F_1^2 and F_2^2 . The structure is similar to figure 3 in Sannino, Pratt & Carillo (2009), based on a three-layer model for the Strait of Gibraltar. Figure 8(b) shows the corresponding contour plot in

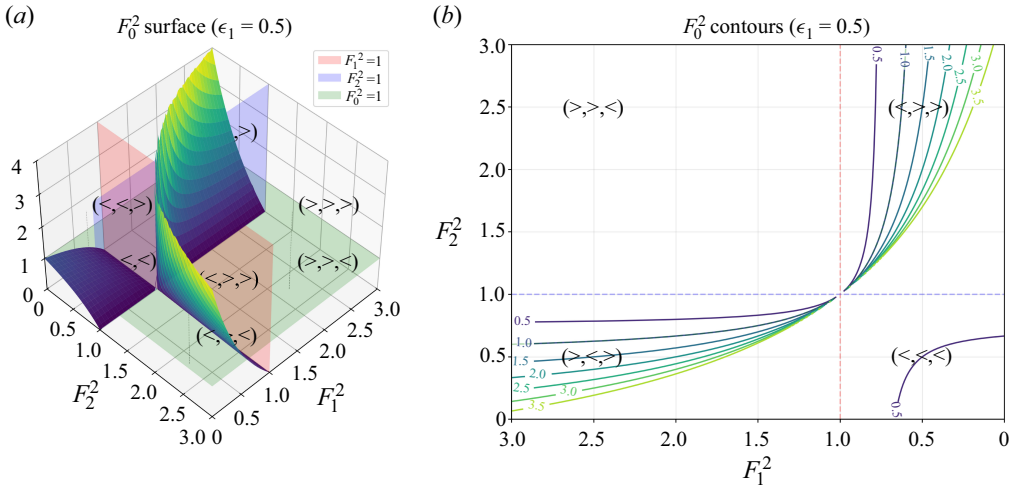


Figure 8. Critical state of the exchange flow. (a) Surface plot of the middle layer Froude number F_0^2 as a function of the upper layer (F_1^2) and lower layer (F_2^2) Froude numbers, satisfying the critical condition given in (3.22). (b) Contour plot of the same critical surface, showing level curves of F_0^2 in the F_1^2 - F_2^2 plane. Regions are labelled using triple inequality notation (e.g. $>, <, >$), indicating whether each layer is supercritical or subcritical relative to the others. These combinations illustrate how individual layers may depart from criticality, while the overall three-layer flow remains in a critical state.

the F_1^2 - F_2^2 plane. Following Sannino *et al.* (2009), Sánchez-Garrido *et al.* (2011), four distinct regimes are identified and labelled using the notation (F_1^2, F_2^2, F_0^2) , where each symbol ($>$ or $<$) denotes whether the corresponding layer is hydraulically supercritical or subcritical relative to the others. These labels illustrate how individual layers may exhibit distinct hydraulic states, even though the overall flow remains critical. For instance, $(>, <, >)$ represents a critical regime where the upper and middle layers are supercritical, while the lower layer is subcritical. Sannino *et al.* (2009) referred to the region marked $(<, <, <)$ as a provisionally subcritical regime, in which all layers are subcritical and wave motions in the upper and lower layers are coupled. They referred to the region $(>, >, <)$ as a provisionally supercritical regime, where the upper and lower layers are supercritical while the middle layer remains subcritical.

In the special case where the middle layer is stagnant ($u_0 = 0$, so that $F_0 = 0$) and either the upper layer or lower layer Froude number reaches unity, i.e. $|F_1| = 1$ or $|F_2| = 1$, a hydraulically controlled region emerges – bounded by $|F_1| = 1$ or $|F_2| = 1$ – in contrast to the localised control points typically found in two-layer flows (Atoufi *et al.* 2023).

To extend this classification beyond the critical state, the characteristics must be related to a composite Froude number G , as introduced in the following section, that reflects the direction of information propagation.

3.3.2. Three-layer hydraulic regimes identifier

The link between characteristics and G can be highlighted by noting $\Pi_{i=1}^4 \lambda_i = a_0/a_4$ combined with (3.20), yielding the identity

$$G = 1 + \frac{(h_1 + h_0 + h_2)}{g_1 g_2 h_1 h_0 h_2} \Pi_{i=1}^4 \lambda_i. \tag{3.23}$$

Although the condition $G = 1$ in a three-layer flow can be used to identify control points, neither (3.21) nor (3.23) can provide insights into whether the three-layer flow is subcritical or supercritical. The reason is that when characteristic speeds of long interfacial waves at both interfaces have imaginary components, the multiplication of all complex characteristic speeds mixes up information propagation associated with each interface and typically gives $G \geq 1$. As only the real components of the characteristics determine the hydraulic regimes (imaginary components quantify the long-wave growth rate, as shown in Appendix A), based on the definition of G , we define a modified composite Froude number

$$\tilde{G} = 1 + \frac{(h_1 + h_0 + h_2)}{g_1 g_2 h_1 h_0 h_2} \text{sign}(\lambda_1^R \lambda_4^R) \prod_{i=1}^4 \lambda_i^R \quad (3.24)$$

that is capable of identifying hydraulic regimes in three-layer flows with complex characteristics. Note that \tilde{G} can be negative, and $\tilde{G} = 1$ at control points, but generally $G \geq \tilde{G}$. When the real components of the four characteristics have the same sign, then $\tilde{G} > 1$, indicating a supercritical regime. Conversely, when $\tilde{G} < 1$, the real components have opposite signs, indicating a subcritical regime. The condition $\tilde{G} = 1$ indicates critical flow since the real component of the characteristic associated with one of the interfaces vanishes.

Thus, \tilde{G} provides a diagnostic based on the direction of information propagation by long interfacial waves, rather than a strict indicator of a hydraulic regime under instability. The modified composite Froude number \tilde{G} is specifically defined to involve only the real parts of the characteristic speeds, avoiding the complications introduced by long-wave instability. As discussed in Appendix D, the real part λ_i^R continues to predict the direction of information propagation by long interfacial waves to leading order, provided the spatiotemporal variation in λ is weak. Therefore, while \tilde{G} should not be interpreted as a strict indicator of a hydraulic regime in the classical sense under instability, it remains a useful diagnostic for identifying regions where long interfacial wave propagation becomes predominantly unidirectional (i.e. supercritical regime), bidirectional (i.e. subcritical regime) or blocked (i.e. critical regime), even when the three-layer flow contains unstable long-wave modes.

3.3.3. Link to turbulence generation

Having established \tilde{G} as an indicator of the hydraulic regime, we now aim to show the link between transition in hydraulic regimes and the turbulence generation and compare predictions from the three-layer model to those of the two-layer model of Atoufi *et al.* (2023). Following Zhu *et al.* (2023), we diagnose the local turbulent kinetic energy (TKE)

$$k'_m \equiv \frac{1}{2} \mathbf{u}'_m \cdot \mathbf{u}'_m, \quad (3.25)$$

computed from the perturbations around a low-pass filtered velocity field

$$\bar{\mathbf{u}}_m(x, y, z, t) \equiv \frac{1}{\Delta L} \int_{-\Delta L/2}^{\Delta L/2} \mathbf{u}(x - s, y, z, t) ds, \quad (3.26)$$

$$\mathbf{u}'_m(x, y, z, t) \equiv \mathbf{u} - \bar{\mathbf{u}}_m. \quad (3.27)$$

We choose the same constant filter width $\Delta L = 10$ to maximise the time- and duct-volume-averaged kinetic energy. Figure 9(a,b,c) shows the normalised TKE averaged over the duct cross-section and figure 9(d,e,f) shows \tilde{G} . For comparison, we also show the modified

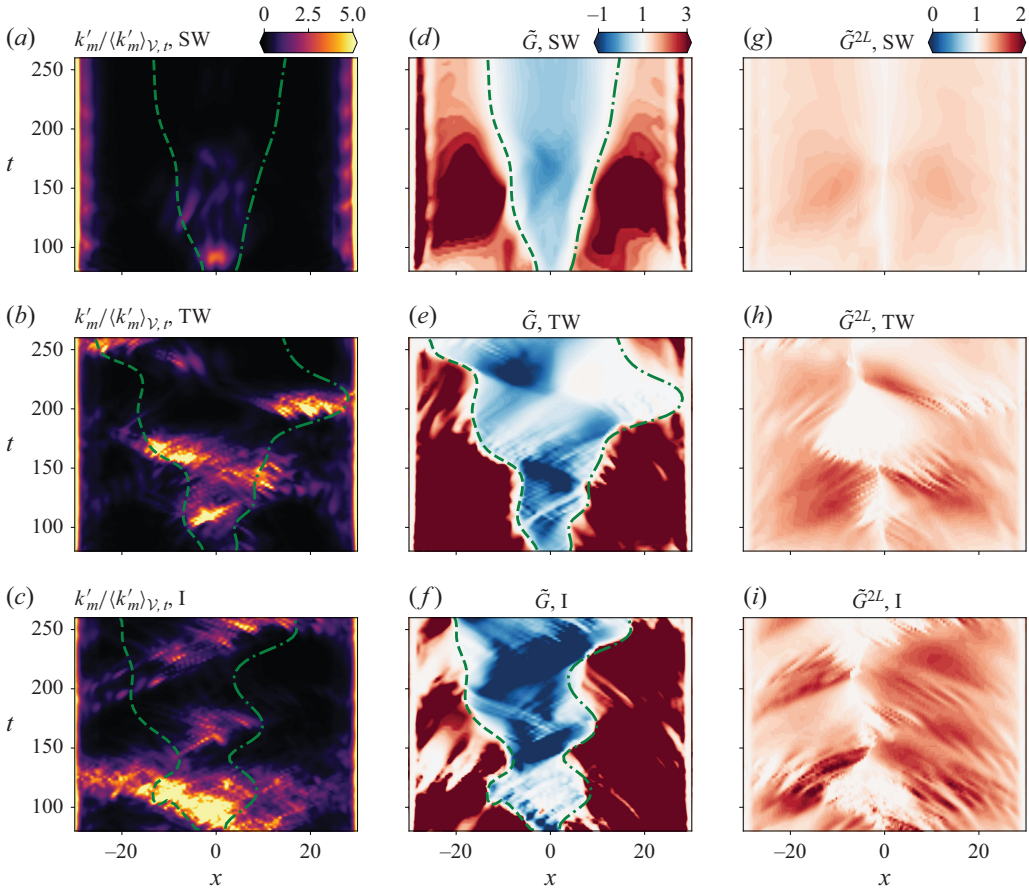


Figure 9. (a–c) Normalised TKE $k'_m/\langle k'_m \rangle_{\mathcal{V},t}$ averaged over the duct cross-section ($\langle \cdot \rangle_{\mathcal{V},t}$ denotes the spatial and temporal average). (d–f) Modified composite Froude number for the three-layer model, \tilde{G} , as defined in (3.24). (g–i) Modified composite Froude number for the two-layer model, \tilde{G}^{2L} , as defined in (3.28). The first row represents the SW case, the second row represents the TW case and the third row represents the I case. The dashed and dashed-dotted green curves are identical to those in figure 3.

composite Froude number for the two-layer hydraulic model,

$$\tilde{G}^{2L} = 1 + \frac{h_1^{2L} + h_2^{2L}}{(\rho_2^{2L} - \rho_1^{2L}) Ri \cos \theta h_1^{2L} h_2^{2L}} \lambda_1^{R^{2L}} \lambda_2^{R^{2L}} \quad (3.28)$$

in figure 9(g,h,i), where the superscripts ‘2L’ denote variables based on the two-layer averaging of Atoufi *et al.* (2023) where the expressions for λ_1 and λ_2 are also given in their (4.10).

As expected, \tilde{G} in figure 9(e) is consistent with figure 6(a,b) for the TW in the sense that it indicates control points as well as the hydraulic regime. In all three cases, the flow experiences a transition from supercritical to subcritical in the central region of the duct where \tilde{G} changes sign. The same auxiliary dashed guidelines (i.e. space–time variation of middle layer height) presented in figures 3 and 6 agree well with $\tilde{G} \approx 1$. We observe a similar agreement between \tilde{G} and the characteristics in the SW and I cases, although their $x-t$ plots are not shown for brevity. Therefore, \tilde{G} identifies hydraulic regimes and control (i.e. flow regularisation when $G = \tilde{G} = 1$ discussed in § 3.4) even in the presence

of mixing and beyond the limit where (3.1) is no longer hyperbolic. Although not shown here, the contours of G have qualitatively similar patterns to \tilde{G} except that $G \geq 1$ in the entire $x-t$ plots for the SW, TW and I cases due to $\lambda^I \neq 0$. Reduced values of G are also observed at the same locations and times where $\tilde{G} < 0$ in figure 9(d,e,f).

In the centre of the duct between the auxiliary curved lines (corresponding to the upper and lower interface locations) in figure 9, the averaged TKE tends to be elevated in the TW and I cases (figure 9b,c), suggesting a correlation between enhanced turbulence and hydraulic regime transitions identified by $\tilde{G} = 1$ (figure 9e,f). In these cases, large values of $\langle k'_m \rangle_{y,z}$ coincide with regions where \tilde{G} transitions from $\tilde{G} > 1$ (supercritical) to $\tilde{G} < 1$ (subcritical), particularly near the duct exits and within central interior regions. In the SW case (figure 9a) this trend is less apparent where the TKE appears visually weak across most of the space-time domain. This reflects a lower level of turbulence rather than a lack of correlation. Nevertheless, localised turbulence is still present, especially near the duct ends and at the central region, which align with changes in \tilde{G} (figure 9d). These observations indicate that, although the spatiotemporal correspondence between TKE and hydraulic regime transitions is strongest in the TW case and more distributed in the I case, it remains qualitatively evident across all three cases. In the I case, small-scale shear instabilities – such as Kelvin–Helmholtz roll-up and breakdown – embedded within the non-hydrostatic pressure gradient force \mathcal{F} also contribute to turbulence generation, in addition to hydraulic effects that are primarily governed by the hydrostatic pressure distribution. As shown in Atoufi *et al.* (2023), the influence of non-hydrostatic pressure is more pronounced in the I case than in the SW and TW cases.

Near both ends, we typically observe short waves forming and breaking leading to turbulence (Zhu *et al.* 2023), suggesting that the flow is unstable to both long- and short-wave instabilities there. As we approach the centre of the duct, the contribution of long waves to the destabilisation of the flow becomes increasingly significant.

For the two-layer model, $\tilde{G}^{2L} = 1$ is more weakly correlated with $\langle k'_m \rangle_{y,z}$. A key result of this paper is that the three-layer \tilde{G} outperforms the two-layer \tilde{G}^{2L} in predicting regions (not only points) where the flow is most unstable and turbulent. Therefore, regime transitions identified by changes in sign of \tilde{G} (i.e. changes in the direction of information propagation) provide excellent predictions of where and when turbulence is generated and strong mixing is expected in stratified exchange flows. The mechanism underlying this agreement is explained in § 4.

3.4. Hydraulic control mechanisms

The characteristics specify the directions of information propagation whereas the control mechanisms specify how information propagates along the characteristics (see § 3.2). Hydraulic control mechanisms are set by some regularity conditions where hydraulic regime transitions occur. To see this link, one can deduce from (3.19) that a solution exists either when $\det \mathbf{A} \neq 0$ or when $\det \mathbf{A} = 0$ with some regularity conditions allowing acceptable solutions. Mathematically, at the control points where $\det \mathbf{A} = 0$ the flow should be regularised by the condition

$$\left\| \text{adj}(\mathbf{A}) \mathbf{B} \frac{\partial s}{\partial x} \right\|_F = 0 \quad (\text{inviscid}) \quad \text{and} \quad \left\| \text{adj}(\mathbf{A}) \tau \mathbf{q} \right\|_F = \mathbf{0} \quad (\text{viscous}), \quad (3.29)$$

where $\|\cdot\|_F$ refers to the Frobenius norm.

3.4.1. Inviscid control mechanism

The inviscid regularity condition in (3.29) implies that

$$\frac{1}{3} \text{adj}(\mathbf{A}) \mathbf{B} \frac{\partial \mathbf{s}}{\partial x} = \begin{pmatrix} g_1 u_1 h_2 (u_0^2 - g_2 h_0 (1 - F_2^2)) \left(\tan \theta - \frac{db}{dx} \right) - g_2 h_2 u_0^2 u_1 \\ g_1 g_2 h_1 h_2 u_0 (1 - F_1^2) + (g_1 h_1)^2 u_0 (1 - F_1^2) \left(\tan \theta - \frac{db}{dx} \right) \\ g_2 u_2 (h_1 u_0^2 - g_1 h_0 h_1 (1 - F_1^2)) - g_1 h_1 u_0^2 u_2 \left(\tan \theta - \frac{db}{dx} \right) \\ g_2 h_1 h_2 u_0^2 - g_1 h_1 (h_2 u_0^2 - g_2 h_0 h_2 (1 - F_2^2)) \left(\tan \theta - \frac{db}{dx} \right) \\ g_1 g_2 h_0 h_1 h_2 (1 - F_2^2) \left(\frac{db}{dx} - \tan \theta \right) - g_1 g_2 h_0 h_1 h_2 (1 - F_1^2) \\ g_1 h_1 h_2 u_0^2 \left(\tan \theta - \frac{db}{dx} \right) - g_2 h_2 (h_1 u_0^2 - g_1 h_0 h_1 (1 - F_1^2)) \end{pmatrix} = \mathbf{0} \tag{3.30}$$

at the control points. We recall that two conditions lead to $\det \mathbf{A} = 0$ and hydraulic control: either (i) $h_0 = 0$ or (ii) $h_0 \neq 0$ but $G = 1$, which further requires $|F_1| = 1$ or $F_2 = 1$ and $F_0 = 0$ (i.e. $u_0 = 0$).

For condition (i), (3.30) enforces $u_0 = 0$. Therefore, both the criticality and inviscid control mechanisms imply that the middle layer velocity vanishes near the jumps. This is consistent with the sudden reduction of $|u_0|$ near the duct ends and close to the control point in figure 4(d). Note that near the duct ends where $|db/dx| \gg \tan \theta$ the flow is also hydraulically controlled as $h_0 \approx 0$ (condition (i)) and $|u_0|$ drops sharply in figure 4(a) to satisfy (3.30). Inside the duct $db/dx = 0$ and $|u_0|$ drops more smoothly.

As stated above, one possibility for $G = 1$ is $F_0 = 0$ and either $|F_1| = 1$ or $|F_2| = 1$ at control points. For condition (ii), (3.30) enforces that $|F_1| = 1$ (critical upper layer) yields $|F_2| = 1$ (critical lower layer), assuming the middle layer has non-zero velocity. However, this implication does not hold in general. For example, when $u_0 = 0$, $F_1 = 0$ and $db/dx = \tan \theta$, all terms in (3.30) vanish, the condition is trivially satisfied and does not constrain F_2 , meaning that both interfaces need not be simultaneously critical. This reflects the fact that when the middle layer is stationary near the control point, the system may exhibit decoupled dynamics at the upper and lower interfaces. Criticality may occur in one interface without enforcing it in the other as also discussed for the special case of pure exchange flow in § 3.2.2. This type of decoupling, particularly relevant for N -layer ($N > 2$) flows with stagnant interior layers, was analysed by Engqvist (1996), who showed that under such conditions, the upper and lower groups of non-stagnant layers can be governed by independent control criteria.

In summary, in addition to flow regularisation, condition (i) implies a thin middle layer and a two-layer exchange flow at both ends. Under condition (ii), when the middle layer is dynamically active, the two oppositely propagating flows at critical velocity ($|F_1| = |F_2| = 1$) cause vigorous mixing ($h_0 \neq 0$) in the middle of the duct (due to symmetry). However, when the middle layer is stationary near the control, the upper and lower interfaces may become decoupled, and criticality may occur independently at one or both interfaces without requiring symmetry.

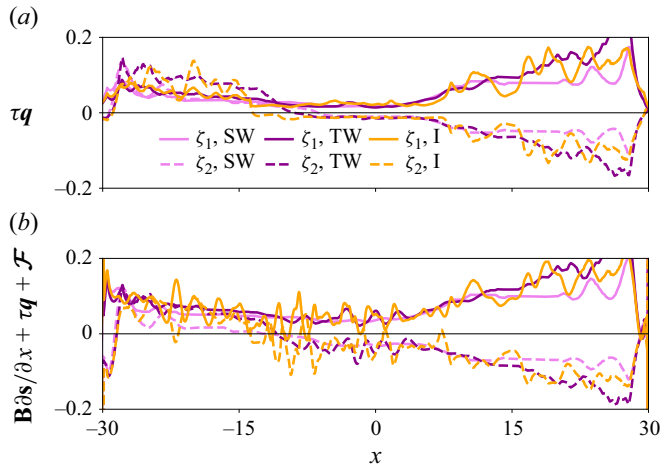


Figure 10. Control mechanisms that dictate how information propagates along the characteristics in the three-layer model at $t = 110$ for all cases: (a) viscous control and (b) the sum of all terms on the right-hand side of (3.1). Solid lines (ζ_1) and dashed lines (ζ_2) represent the first and second momentum equations (first and second rows, respectively) in (3.1). The lines are shown for the SW, TW and I cases in light purple, dark purple and yellow, respectively.

3.4.2. Viscous control mechanism

The effects of viscosity are of particular importance as they allow for enhanced dissipation in hydraulic jumps, and are connected to turbulent mixing when the hydraulic regime transitions from supercritical to subcritical.

In figure 10 we computed the viscous control term τq (non-zero elements) to verify that it indeed dominates the right-hand side of (3.1) in the central region of the duct where $\tilde{G} = 1$ (at a representative time $t = 110$). Panels (a,b) present the viscous terms and the sum of all the terms on the right of (3.1), respectively, and show that the viscous term dominates, particularly in the central region of the duct. For the wave cases, the viscous control mechanism, represented by τq , becomes $\lesssim 0.01$ in the centre of the duct to regulate the flow (§ 3.4) at the internal control points. The summation of these terms also reaches a local minimum near the control points. Note that the viscous term and sum also become approximately zero at the duct ends, where geometric control is established. The regularity condition is therefore satisfied for all control mechanisms at the duct ends.

The large contribution of viscous terms τq suggests that information propagation along the interfaces is predominantly modulated by viscous effects (see § 3.2). By comparing figures 7 and 10, we observe that along the upper interface, viscous damping occurs on the left side of the duct, where $\zeta_1 = (\tau q)_1 > 0$ (with the subscript 1 indicating the term from the velocity difference equation between the middle and upper layers, i.e. the first row of (3.1)) and $\lambda_1^R > 0$, indicating viscous dissipation. On the right side of the duct, where $\zeta_1 = (\tau q)_1 < 0$ and $\lambda_1^R < 0$, the viscous forcing reinforces the direction of wave propagation. This results in a local amplification of interfacial waves due to enhanced mixing with the lower layer. Similarly, along the lower interface, viscous damping is observed on the right side, where $\zeta_2 = (\tau q)_2 < 0$ and $\lambda_3^R > 0$, again indicating dissipation. Here, the subscript 2 refers to the velocity difference equation between the lower and middle layers (the second row of (3.1)). On the left side, where $(\tau q)_2 > 0$ and $\lambda_3^R < 0$, viscous forcing contributes constructively to wave propagation, leading to amplification through mixing with the upper layer.

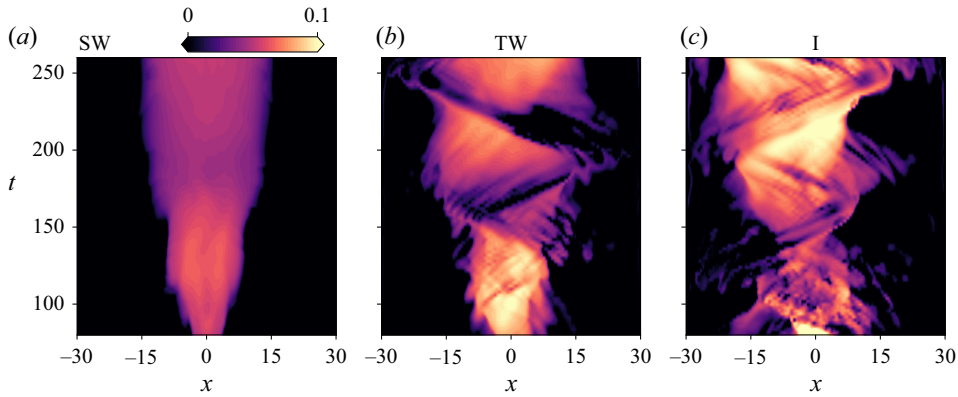


Figure 11. Space–time plots of the long-wave growth rate simultaneity, $\lambda_2^I \lambda_4^I$, for (a) SW, (b) TW and (c) I cases.

Thus far our focus has been primarily on the physical significance of characteristics in hydraulic control and regime transition embedded in λ^R . Although we showed a strong correlation between turbulence generation and λ^R , to reveal their true connections, we need to consider instabilities through the imaginary part λ^I , as we do next.

4. Three-layer instabilities and resonances

4.1. Simultaneous amplification in the growth rate of the long waves

In [Appendix A](#) we show that the characteristics λ obtained from (3.9) are identical to the phase speed of linear long waves (with wavenumber $k \ll 1$) propagating in a background three-layer flow. In § 3 we discussed the essential roles of the λ^R to predict the control points and hydraulic regimes, which are closely associated with the laminar–turbulence transition in SID flows. Nonetheless, a consequence of the equality of the λ and phase velocity of the linear long wave is that $\lambda^I > 0$ represents their growth rate. Therefore, it is reasonable to anticipate that the space–time distribution of the imaginary components presented in [figure 6\(c,d\)](#) may be a proxy to the TKE shown in [figure 9\(a,b,c\)](#). However, upon comparing their distributions, it is evident that this is not the case. In [figure 6\(c\)](#) the imaginary components of characteristics of the upper interface λ_1^I exhibit a positive growth rate in the right and middle regions of the duct (separated by the auxiliary lines in [figure 6\(c\)](#)), while λ_3^I associated with the lower interface displays a positive imaginary component in the middle and right regions of the duct. They do not match with the TKE distribution in [figure 9\(b\)](#), which shows the largest TKE in the middle of the duct. However, when multiplying $\lambda_2^I(x, t)$ and $\lambda_4^I(x, t)$ to diagnose whether both interfaces simultaneously grow or not, we expect $\lambda_2^I \lambda_4^I(x, t)$ to agree with the TKE distribution, which will be discussed next.

We show in [figure 11](#) the spatiotemporal distribution of the simultaneous amplification $\lambda_2^I \lambda_4^I(x, t)$, reflecting the growth of long waves at the upper and lower interfaces in the SW (a), TW (b) and I (c) regimes. In all cases, the central region of the duct exhibits the largest $\lambda_2^I \lambda_4^I$ and the amplitude growth of the long waves (indicated by light colours) occurs at the same rate, even though λ_2^I and λ_4^I attain their highest values in regions corresponding to dark contours. A qualitative correspondence can be observed between regions of elevated $\lambda_2^I \lambda_4^I$ in [figure 11](#) and enhanced TKE in [figure 9\(a–c\)](#), particularly in the TW and I cases. In these cases, simultaneous growth in the amplitude of long

interfacial waves is followed (with some time delay) by the emergence of turbulence, supporting a linkage between long-wave amplification and the eventual onset of smaller-scale generation through nonlinear energy transfer and wave breaking. In the SW case, however, this correspondence is less clear due to weaker turbulence levels. Differences across the cases may also arise from the emergence of short-wave instabilities, which are addressed in the following section.

Further understanding of the simultaneous growth in a three-layer exchange flow can be gained by building-block decomposition of the inviscid homogeneous limit of (3.1), which will be discussed next.

4.2. *Instability arising from long waves: isolated-layer perturbation*

Although λ^I represents the growth rate of the linear long waves (see Appendix A), the origin of this imaginary component in the characteristics remains unclear. In two-layer flows the presence of λ^I is linked to the cross-interface shear exceeding a certain threshold, which induces the appearance of the imaginary component (Long 1956; Dalziel 1991; Lawrence 1993; Atoufi *et al.* 2023). However, for three-layer flows, identifying a single threshold is not feasible due to the increased complexity and number of variables involved. The resulting analytical expression for the characteristics, as given in (3.12), is too lengthy and intricate to define a simple criterion. Therefore, we seek an alternative approach that relates λ^I to interactions between the layers.

We focus on the free propagation of the long waves described by the inviscid homogeneous limit of (3.1) where $\mathbf{C}(\partial\mathbf{q}/\partial t) + \mathbf{A}(\partial\mathbf{q}/\partial x) = \mathbf{0}$ and we decompose the coefficient matrices as

$$\begin{aligned} \mathbf{A} &= \left(\mathbf{A}_1 + \frac{\mathbf{A}_0}{2} \right) + \left(\mathbf{A}_2 + \frac{\mathbf{A}_0}{2} \right), \\ \mathbf{C} &= \left(\mathbf{C}_1 + \frac{\mathbf{C}_0}{2} \right) + \left(\mathbf{C}_2 + \frac{\mathbf{C}_0}{2} \right), \end{aligned} \tag{4.1}$$

such that

$$\mathbf{A}_i = \begin{pmatrix} -u_1\delta_{i1} & u_0\delta_{i0} & 0 & 0 & g_1 & g_1 \\ 0 & -u_0\delta_{i0} & u_2\delta_{i2} & 0 & 0 & g_2 \\ -h_1 & h_0 & 0 & -u_1\delta_{i1} & u_0\delta_{i0} & 0 \\ 0 & -h_0 & h_2 & 0 & -u_0\delta_{i0} & u_2\delta_{i2} \\ 0 & 0 & 0 & 1 & 1 & 1 \\ h_1 & h_0 & h_2 & u_1\delta_{i1} & u_0\delta_{i0} & u_2\delta_{i2} \end{pmatrix}, \tag{4.2}$$

$$\mathbf{C}_i = \begin{pmatrix} -\delta_{i1} & \delta_{i0} & 0 & 0 & 0 & 0 \\ 0 & -\delta_{i0} & \delta_{i2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{4.3}$$

where $i = 0, 1, 2$ and δ_{ij} denotes the Kronecker delta function. To facilitate theory, we introduce the coefficient matrix pairs

$$\mathbf{W} = (\mathbf{A}_1, \mathbf{C}_1), \quad \mathbf{Y} = (\mathbf{A}_2, \mathbf{C}_2), \quad \mathbf{Z} = (\mathbf{A}_0, \mathbf{C}_0). \tag{4.4}$$

The first matrix pair, \mathbf{W} , represents an isolated upper layer moving over stationary middle and lower layers, as if these two layers act as a hypothetical topography that restricts the

frictionless motion of the upper layer. Notably, u_0 and u_2 from \mathbf{A} and the corresponding columns in \mathbf{C} do not appear in \mathbf{W} , reflecting the isolation of the upper layer. Similarly, the second matrix pair, \mathbf{Y} , denotes the isolated lower layer moving without friction beneath the stationary middle and upper layers, which act as a topography restricting its motion. Here, u_0 and u_1 from \mathbf{A} and their corresponding columns in \mathbf{C} are absent from \mathbf{Y} . Finally, the third matrix pair, \mathbf{Z} , represents an isolated middle layer moving without friction between the ‘topography’ provided by the stationary upper and lower layers.

The generalised eigenvalue problem associated with each matrix pair has only real solutions of the form:

$$\lambda(\mathbf{W}) = u_1 \pm \sqrt{g_1 h_1}, \tag{4.5}$$

$$\lambda(\mathbf{Z}) = \left(u_0 - \sqrt{\frac{g_1 g_2}{g_1 + g_2} h_0} \right) \left(u_0 + \sqrt{\frac{g_1 g_2}{g_1 + g_2} h_0} \right), \tag{4.6}$$

$$\lambda(\mathbf{Y}) = u_2 \pm \sqrt{g_2 h_2}. \tag{4.7}$$

Now consider that the isolated upper layer matrix pair \mathbf{W} is perturbed by $\mathbf{Z}/2$ through the interaction between the upper and middle layers. Similarly, we allow $\mathbf{Z}/2$ to perturb the isolated lower layer, accounting for the interaction between the lower and middle layers. We denote these perturbed layers with the following matrix pairs: $\tilde{\mathbf{W}} = \mathbf{W} + \mathbf{Z}/2$, $\tilde{\mathbf{Y}} = \mathbf{Y} + \mathbf{Z}/2$, and we note that $\tilde{\mathbf{X}} = (\mathbf{A}, \mathbf{C}) = \tilde{\mathbf{W}} + \tilde{\mathbf{Y}}$. We refer to $\tilde{\mathbf{W}}$ as the perturbed upper layer matrix pair and $\tilde{\mathbf{Y}}$ as the perturbed lower layer matrix pair, noting that $\tilde{\mathbf{X}}$ yields the original three-layer exchange flow in the inviscid homogeneous limit.

Since the flow is stably stratified, with $g_1, g_2 > 0$, all eigenvalues in (4.5) are real. However, the perturbed matrix pairs introduce imaginary components in their characteristics, such that $\lambda^l(\tilde{\mathbf{W}}) \neq 0$, $\lambda^l(\tilde{\mathbf{Y}}) \neq 0$ in general. Therefore, attributing the criterion $\lambda^l(\tilde{\mathbf{X}}) > 0$ to the onset of long-wave instabilities (see Appendix A) reveals that the instability arises exclusively from $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{Y}}$, the perturbed upper and lower isolated layers influenced by the middle layer. We note that $\lambda(\mathbf{W} + \mathbf{Y}) \in \mathbb{R}$ meaning that perturbing \mathbf{W} by \mathbf{Y} alone (and *vice versa*) does not introduce imaginary components in the characteristics. This implies that the origin of the imaginary components results specifically from the interaction of the isolated upper and lower layers with the middle layer.

The matrix pairs $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{Y}}$ represent a building-block decomposition, illustrating that the full system $\tilde{\mathbf{X}}$ comprises two interacting subsystems: the perturbed upper layer ($\tilde{\mathbf{W}}$) and the perturbed lower layer ($\tilde{\mathbf{Y}}$).

In a three-layer system, three types of resonance are possible: (i) resonance between interfacial waves at the upper interface alone, (ii) resonance at the lower interface alone and (iii) resonance due to interactions between waves propagating on both interfaces. The decomposition into $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{Y}}$ also offers a framework for diagnosing all three cases by isolating contributions from each layer and their coupling.

To illustrate this, we consider the special case of pure exchange flow discussed in § 3.2.2. In this case, $\|\tilde{\mathbf{W}}\|_F = \|\tilde{\mathbf{Y}}\|_F$, leading to $\lambda_{1,2} = \lambda(\tilde{\mathbf{Y}})$ and $\lambda_{3,4} = \lambda(\tilde{\mathbf{W}})$. Recall that when $F_1^2 > F_{\Delta+}^2$, the phase speed of long waves associated with a single interface acquire imaginary components due to increased shear. This can be understood as the result of interaction between the isolated upper and middle layers, which generates sufficient shear, reflected in $\lambda^l(\tilde{\mathbf{W}})$, to destabilise the upper interfacial waves. Similarly, interaction between the isolated lower and middle layers can produce shear strong enough to drive instability, captured by $\lambda^l(\tilde{\mathbf{Y}})$. Thus, in the absence of coupling between interfacial waves,

the imaginary parts $\lambda^I(\tilde{\mathbf{W}})$ and $\lambda^I(\tilde{\mathbf{Y}})$ individually capture instabilities arising at the upper and lower interfaces, respectively.

In more general cases where the middle layer is non-stagnant and the upper and lower layers are not symmetric, so that $\lambda_{1,2} \neq \lambda(\tilde{\mathbf{Y}})$ and $\lambda_{3,4} \neq \lambda(\tilde{\mathbf{W}})$, the characteristics of the individual building blocks still provide a useful diagnostic. By comparing $\lambda^I(\tilde{\mathbf{W}})$ and $\lambda^I(\tilde{\mathbf{Y}})$ with the full system $\lambda^I(\tilde{\mathbf{X}})$, one can assess whether the dominant instability arises at the upper interface (i), the lower interface (ii) or from cross-interface coupling (iii).

In this study we primarily focus on the third type of resonance – interfacial interactions across the layers. In the next section we examine the instability mechanisms that emerge from resonance between long interfacial waves propagating on different interfaces.

4.3. Spectral gap and long-wave resonance indicator

In the preceding section we demonstrated that a pathway for the long-wave amplitude growth $\lambda^I(\tilde{\mathbf{X}})$ originates from the growing waves in perturbed upper and lower layers denoted by $\lambda^I(\tilde{\mathbf{Y}})$ and $\lambda^I(\tilde{\mathbf{W}})$, respectively. In a three-layer flow, the long waves associated with these perturbed lower and upper layers can resonate, to amplify the imaginary component of the characteristics of the full three-layer system $\lambda(\tilde{\mathbf{X}})$.

4.3.1. Diagnosing long-wave resonance across interfaces from DNS

In matrix perturbation theory (Stewart & Sun 1990), the term spectral gap is often used to describe the difference between the spectra of two matrix pairs. Given a matrix pair (\mathbf{A}, \mathbf{C}) , where we consider the generalised eigenvalue problem $\mathbf{A}\mathbf{v} = \lambda\mathbf{C}\mathbf{v}$, a perturbation of the matrix pair results in a perturbed generalised eigenvalue problem $(\mathbf{A} + \Delta\mathbf{A}, \mathbf{C} + \Delta\mathbf{C})$. The spectral gap in this context quantifies the variation in eigenvalues caused by the perturbation of both matrices in the pair. This measure quantifies the effects of small changes to both matrices on the generalised eigenvalues and, consequently, the stability and dynamics of the system. In the context of perturbed upper and lower layers (represented by matrix pairs $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{Y}}$, respectively), the spectral gap measures how the spectra of these perturbed layers deviate from each other and from the isolated middle layer represented by \mathbf{Z} . By examining the spectral gap, we gain insight into how resonant interactions between perturbed layers influence the characteristics of the full three-layer system. Further explanations and mathematical descriptions for the spectral gap in the context of perturbed layers are provided in [Appendix B](#).

To quantify such resonant interactions in DNS data, we measure the relative spectral gap between (characteristics of) perturbed upper and lower layers with respect to the (characteristics of) isolated middle layer as

$$\mathcal{R}(\mathbf{W}, \mathbf{Y}) = \max_i \max_j \frac{|\Delta\lambda_i^{\tilde{\mathbf{W}}}| \sqrt{1 + |\lambda_j(\tilde{\mathbf{Y}})|^2} - |\Delta\lambda_j^{\tilde{\mathbf{Y}}}| \sqrt{1 + |\lambda_i(\tilde{\mathbf{W}})|^2}}{|\Delta\lambda_i^{\tilde{\mathbf{W}}}| \sqrt{1 + |\lambda_j(\tilde{\mathbf{Y}})|^2} + |\Delta\lambda_j^{\tilde{\mathbf{Y}}}| \sqrt{1 + |\lambda_i(\tilde{\mathbf{W}})|^2}}, \quad (4.8)$$

where $\Delta\lambda_i^{\tilde{\mathbf{W}}} = [\lambda_i^R(\tilde{\mathbf{W}}) - \lambda(\mathbf{Z})] + I\lambda_i^I(\tilde{\mathbf{W}})$ and $\Delta\lambda_j^{\tilde{\mathbf{Y}}} = [\lambda_j^R(\tilde{\mathbf{Y}}) - \lambda(\mathbf{Z})] + I\lambda_j^I(\tilde{\mathbf{Y}})$. The superscripts R and I refer to real and imaginary components of the corresponding characteristics and $i, j = 1, \dots, 4$ are indices counting individual characteristics. Note that $\mathcal{R}(\mathbf{W}, \mathbf{Y}) = \mathcal{R}(\mathbf{Y}, \mathbf{W})$. The reader is referred to [Appendix B.2](#) to see a further description and mathematical derivation of (4.8).

Equation (4.8) yields that under the condition where $\lambda^R(\tilde{\mathbf{W}}) = \lambda^R(\tilde{\mathbf{Y}}) = \lambda(\mathbf{Z})$ and $\lambda^I(\tilde{\mathbf{Y}}) = \lambda^I(\tilde{\mathbf{W}}) > 0$, then $\mathcal{R}(\mathbf{W}, \mathbf{Y}) = 0$, which implies that long waves can resonantly

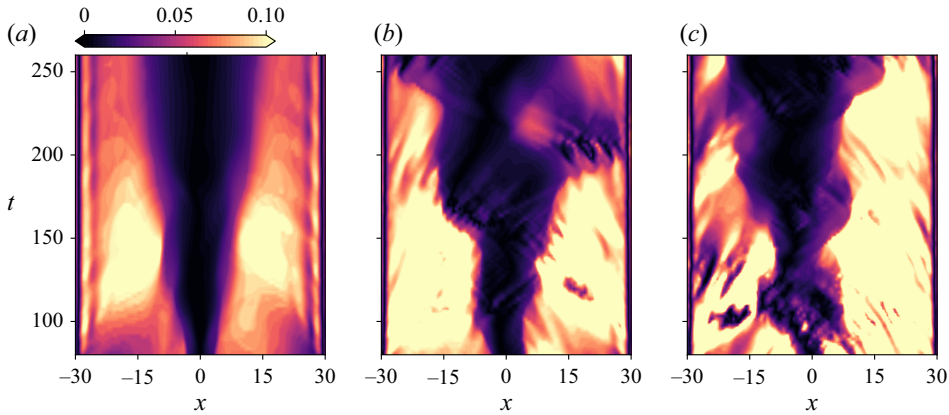


Figure 12. Contours of the long-wave resonance indicator \mathcal{R} plotted in space–time (x, t) for the SW (a), TW (b) and I (c) cases. The dark regions indicate the minimum spectral gap, as defined in (4.8), between the perturbed upper and lower layers, derived from the building-block matrix decomposition of the three-layer system into perturbed upper and lower layer subsystems in (4.1).

interact since they are phase locked and have an equal growth rate. Therefore, $\mathcal{R} \rightarrow 0$ (note that $0 \leq \mathcal{R} \leq 1$) is a condition for near resonance, whereas when $\mathcal{R} \rightarrow 1$, long waves are not phase locked and do not resonate. Note that for the special case of pure exchange flow discussed in § 3.2.2, the $F_1^2 = F_{\Delta+}^2$ condition that defines the neutral stability curve for long waves also leads to $\lambda(\tilde{\mathbf{Y}}) = \lambda(\tilde{\mathbf{W}})$, which corresponds to the resonance condition $\mathcal{R} = 0$ (since $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$). Equivalently, $\mathcal{R} = 0$ marks the neutral stability of the long waves.

This equivalence highlights a connection between the resonance mechanism and the onset of long-wave instability in pure exchange flows. For more general three-layer exchange flows, where the neutral stability curve cannot be expressed analytically, we expect the resonance measure \mathcal{R} to continue to serve as a useful diagnostic for identifying resonant interactions between interfacial waves. To confirm this, we apply the \mathcal{R} measure to DNS data in the following section.

4.3.2. Application to the data

The contours of the long-wave resonance indicator \mathcal{R} are shown in figure 12 for the SW, TW and I cases. All cases qualitatively correspond to the \tilde{G} contours in figure 9(d–f), particularly in the sense that the regime transitions from a supercritical hydraulic regime ($\tilde{G} > 1$) to a subcritical regime ($\tilde{G} < 1$) when $\mathcal{R} \approx 0$. This similarity is noteworthy and not self-evident as the characteristics of the perturbed upper layer and those of the upper interface in the original three-layer flow are different by definition, i.e. $\lambda(\tilde{\mathbf{W}}) \neq \lambda_{1,2}(\hat{\mathbf{X}})$, and, similarly, $\lambda(\tilde{\mathbf{Y}}) \neq \lambda_{3,4}(\hat{\mathbf{X}})$ in general. Therefore, one would expect that $\lambda(\tilde{\mathbf{Y}}) + \lambda(\tilde{\mathbf{W}}) \neq \lambda(\hat{\mathbf{X}})$. The striking resemblance between the spatiotemporal distributions of \mathcal{R} and \tilde{G} suggests that the building-block decomposition introduced in § 4.2 provides valuable insight not only into the pathway for growth of interfacial waves but also their interactions and three-layer hydraulics. Moreover, it suggests that the hydraulic regime transition, driven by changes in the direction of information propagation along the interfaces, is also accompanied by long-wave resonance.

This long–long-wave resonance will modify the base state (mean flow) by the long waves themselves (figure 2). Such changes in the base state may lead to the onset of

short-wave instabilities and result in the TKE generation at later times, as also found in Zhu *et al.* (2024). For all cases, the contours of \mathcal{R} are consistent with the TKE contours presented in figure 9 in the sense that strong TKE events are linked with interfacial wave resonances (i.e. the strong TKE occurs generally in places with the smallest \mathcal{R}). A more detailed comparison between instantaneous profiles of TKE and \mathcal{R} is presented in Appendix C.

To properly establish the link between long-wave resonance and TKE generation, we also identify the resonance between long-wave and short-wave packets, eventually leading to the generation of small scales, which will be discussed next.

4.4. Indication of resonance between long and short waves

In this section we investigate the resonant interaction between long and short waves and provide evidence for the link between these interactions and the TKE. We begin by examining the connection between long- and short-wave instability in hydraulically controlled flow (§ 4.4.1). We then identify the conditions under which resonance between these waves can occur (§ 4.4.2). In § 4.4.3 we explore how such resonance contributes to turbulence generation, extending the analysis beyond linear wave stability. Finally, we apply the resulting diagnostics and criteria to DNS data in § 4.4.4.

4.4.1. Linking long-wave and short-wave instability: hydraulically controlled instability

In the previous section we examined the link between internal hydraulics and the instability of long waves. We now show that the stability of the entire wave spectrum – including short and intermediate waves – is likewise governed by the internal hydraulic state, characterised by the layer Froude numbers. This mechanism, which we refer to as hydraulically controlled instability, highlights that although the system is a stratified shear flow both the shear and stratification are prescribed by hydraulic conditions, which in turn control stability across all wavenumbers.

To illustrate this, we again consider the pure exchange flow case discussed in § 3.2.2, which simplifies the analysis. The imaginary components of the phase speeds of all waves in the spectrum, derived from the linear theory developed in Appendix A, are shown in figure 13. The top row corresponds to the bifurcation Froude number $F_{\Delta-}^2$ and the bottom row to the marginal-stability Froude number $F_{\Delta+}^2$. As the Froude number increases from $F_{\Delta-}^2$ to $F_{\Delta+}^2$, the wave instability structure changes qualitatively across all wavenumbers – not just in the long-wave limit.

Figure 13(a,b) shows that when $F_1^2 = F_{\Delta-}^2$, unstable modes span a broad wavenumber range – from long to short waves – up to $k \lesssim 10$, particularly in three-layer flows with a thick middle layer ($r \geq 1$), where r denotes the ratio of middle to upper layer thickness. Moreover, the phase speeds c_1 and c_4 exhibit similar imaginary components across wavenumbers, indicating that interfacial waves at both interfaces grow at nearly the same rate. In contrast, when $F_1^2 = F_{\Delta+}^2$, corresponding to marginal long-wave stability, instability is confined to short waves, which persist across a wide range of r , as shown in figure 13(c,d). In this case, the c_1^I and c_4^I differ significantly, both in the values of r that support instability and in the wavenumber ranges associated with peak c_1^I and c_4^I , reflecting a decoupling of instability mechanisms across the layers.

These results suggest that the layer Froude numbers dictate both the onset and extent of instability across the wave spectrum. The observed similarity in the c_1^I and c_4^I in the $F_1^2 = F_{\Delta-}^2$ case between long- and short-wave behaviour raises the question of whether these waves can interact resonantly – a possibility we now explore.

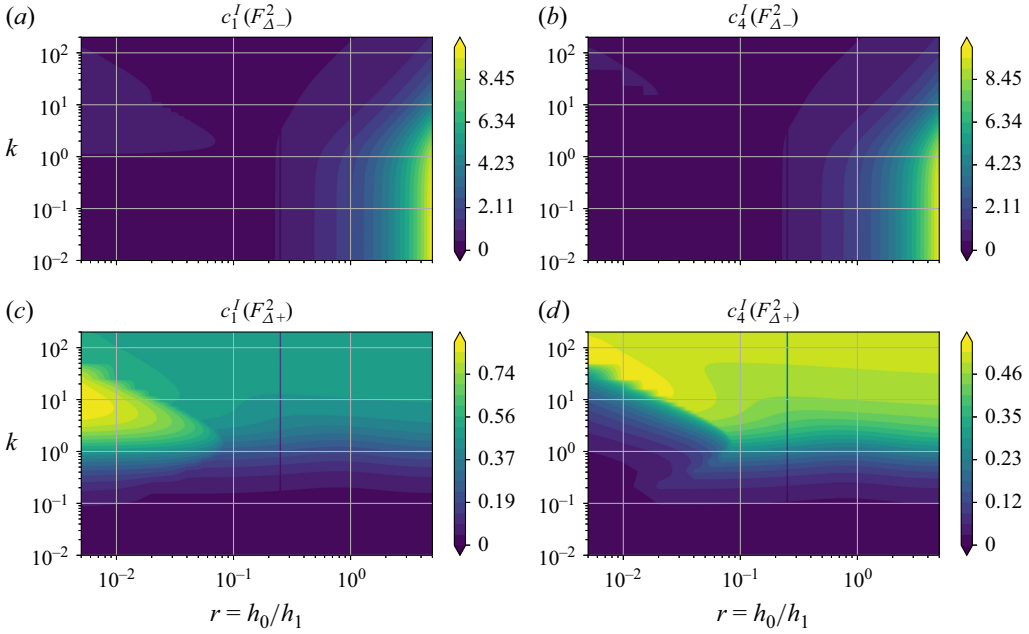


Figure 13. Instability map from contours of the imaginary component of the phase speed, c^I , for linear disturbance waves in a three-layer pure exchange flow with a stagnant middle layer introduced in § 3.2.2. Both the bifurcation Froude number case $F_1^2 = F_{\Delta-}^2$ and the marginal-stability Froude number case $F_1^2 = F_{\Delta+}^2$ defined in (3.16) are shown, isolating the dependence of wave amplification on the layer depth ratio r and wavenumber k . For $F_{\Delta-}^2$, the interfacial modes c_1 and c_4 exhibit similar values of c^I , indicating comparable growth rates at both interfaces. In contrast, for $F_{\Delta+}^2$, the growth characteristics differ between interfaces, with distinct ranges of r and k contributing to instability.

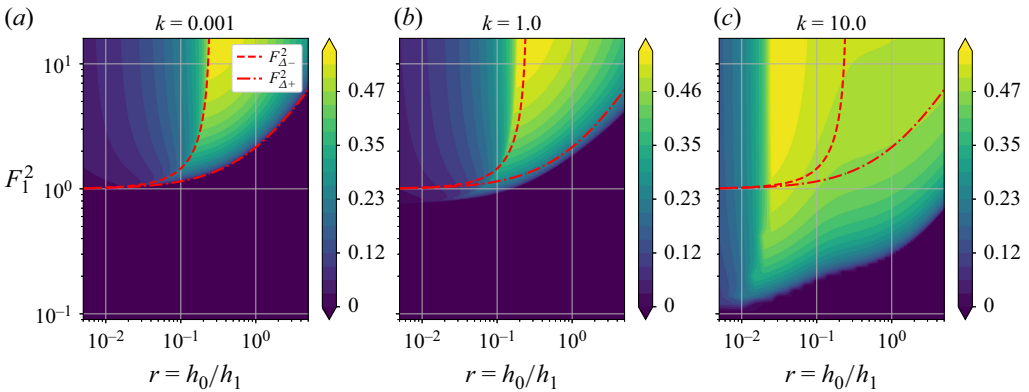


Figure 14. Imaginary component of the phase speed of the lower interfacial wave, c_4^I , for pure exchange flow with a stagnant middle layer introduced in § 3.2.2 at selected wavenumbers $k = 0.001$ (a), 1 (b) and 10 (c), plotted as a function of layer thickness ratio r and Froude number F_1^2 .

4.4.2. Resonance between long and short waves

To assess the potential for resonance between long and short waves, figure 14 shows c_4^I for the pure exchange flow at selected wavenumbers ($k = 0.001$, 1 and 10), across a range of layer thickness ratios r and Froude numbers. At $k = 0.001$ shown in figure 14(a), c_4^I closely matches its value in the long-wave limit ($k \rightarrow 0$) shown previously in

figure 13(b), confirming that the solution represents a long-wave mode. As can be seen from figure 14(a–c), for $k < 1$, the variation in phase speed as the wavenumber changes is relatively weak, while for $k > 1$, the phase speed continues to change significantly. This suggests that dispersion is strongest around $k = 1$, where small differences in wavenumber lead to substantial differences in propagation speed. This behaviour implies that wave packets centred near $k = 1$ are highly dispersive – that is, their constituent wavenumbers propagate at different phase speeds and, consequently, different group velocities. The implications of this for resonance and energy exchange will be explored in the following section. Additionally, increasing k broadens the region of instability as shown in figure 14(a–c), extending to lower Froude numbers ($F_1^2 < F_{\Delta+}^2$) and a wider range of r , indicating that short-wave instability becomes more prominent where $F_1^2 < F_{\Delta+}^2$.

For systematic identification of resonance between long and short waves, let us recall that a packet of linear waves with phase velocity c propagates at the group velocity,

$$c_g = \frac{\partial(ck)}{\partial k} = c + k \frac{\partial c}{\partial k} = c + kd_g, \tag{4.9}$$

where $d_g = \partial c / \partial k$ measures the wave dispersivity. The kinetic energy is carried by the group velocity c_g as the wave packet propagates, and c is the phase velocity, whose mathematical expression is defined in Appendix A (from (A10)). Of particular interest is a packet of short waves with a group velocity close to the phase speed of the long waves (i.e. $c_g = \lambda$), since long and short waves can then resonate (Ma 1981).

In general, $c \in \mathbb{C}$, so c_g also contains real and imaginary parts. The real part c_g^R represents the propagation speed of the wave envelope, while the imaginary part c_g^I corresponds to its temporal growth. Resonance between wave packets and long interfacial waves is most likely when both $c_g^R \approx \lambda^R$ and $c_g^I \approx \lambda^I$.

We utilise the general form of the spectral gap, as defined in (B13), to assess the relative group velocity of short wave packets compared with the phase speed of long waves at both the lower and upper interfaces. For a given wavenumber k_0 , we interpret the group velocity $c_g|_{(k=k_0)} \in \mathbb{C}$ as the generalised eigenvalue of a perturbed matrix pair involving the three-layer matrix pair $\widehat{\mathbf{X}} = (\mathbf{A}, \mathbf{C})$, expressed as

$$c_g|_{(k=k_0)} = \lambda(\widehat{\mathbf{X}} + \mathbf{M}''), \tag{4.10}$$

where \mathbf{M}'' is a perturbation matrix pair that accounts for the finite-wavenumber modulation of the wave packet. We recall that $\lambda(\widehat{\mathbf{X}})$ yields the phase speed and growth rate of long interfacial waves in the three-layer flow, while $\lambda(\widehat{\mathbf{X}} + \mathbf{M}'')$ provides the corresponding quantities for the amplitude envelope of a wave packet centred around k_0 . Using this interpretation, we apply the spectral gap definition from (B13) to compute

$$\mathcal{S} \equiv \mathcal{S}_{\widehat{\mathbf{X}}}(\mathbf{M}''), \tag{4.11}$$

i.e. the spectral gap between long waves and short wave packets, by substituting $\mathbf{M} = \widehat{\mathbf{X}}$ and $\mathbf{M}' = \mathbf{M}''$ in (B13).

The spectral gap \mathcal{S} thus provides a diagnostic criterion for identifying near-resonant interactions between short-wave packets and long interfacial waves. Unlike the dispersion relation, which describes the behaviour of individual modes, \mathcal{S} quantifies the extent to which a short-wave packet can propagate at the same speed as a long wave. When this condition is met, efficient energy transfer from the long-wave to the short-wave packet becomes possible. In the next section we examine how this resonance mechanism contributes to turbulence generation, particularly through the amplification of short-wave disturbances in regions where \mathcal{S} is small.

4.4.3. The role of resonance in turbulence generation

The product $c_2^I c_4^I$, shown in [figure 17\(c\)](#) for the TW and I cases, highlights regions in (x, k) space where both interfacial modes are simultaneously unstable. These regions of modal amplification are especially pronounced for $k \lesssim 1$ and spatially coincide with zones of elevated TKE in [figure 9](#). Although the peak values of c_2^I and c_4^I do not occur at the same x for all wavenumbers, their product effectively identifies regions of overlap where both modes contribute to instability. This spatial alignment suggests that energy amplification is not solely attributable to isolated modal growth, but may be enhanced by interactions between wave modes on both interfaces. In particular, simultaneous instability at the upper and lower interfaces permits resonant coupling between long interfacial waves and dispersive short-wave packets, enabling efficient energy transfer across scales.

To understand the conditions that allow this energy transfer, we consider the limitations of the dispersion relation alone. While the dispersion relation derived in [Appendix A](#) plays a central role in identifying both long- and short-wave instabilities, it does not capture the interactions between modes at different scales. For instance, [figure 17](#) shows that unstable modes exist on both interfaces over a broad wavenumber range. However, the dispersion relation does not indicate whether these modes are phase locked, whether their group velocities coincide or whether energy can be transferred between them.

This motivates the introduction of the spectral gap \mathcal{S} , which quantifies the proximity between the group velocity of a short-wave packet and the phase speed of a long wave. When $\mathcal{S} \rightarrow 0$, a near-resonant condition is established: the short-wave packet is not only unstable but can also efficiently extract energy from long energetic waves. This leads to resonance-driven amplification of short-wave disturbances, even in parameter regimes where their modal instability alone may not explain the observed turbulence.

4.4.4. Application to the data

To apply the theoretical framework developed above, we must select a representative wavenumber k_0 around which to evaluate the spectral gap \mathcal{S} . This choice is guided by both dispersivity analysis and prior evidence of potential resonance. In particular, from the discussion in [§ 4.4.2](#), we have already seen that short-wave packets centred at $k_0 = 1$ exhibit phase speeds comparable to those of long interfacial waves, suggesting conditions favourable for resonance.

This is further supported by the dispersivity analysis shown in [figure 18](#), which reveals that the wave dispersivity $|d_g|$ peaks at $k_0 = 1$, indicating the strongest dispersion and the wavenumber, therefore, is effective for coherent wave packet formation that carries energy through the flow. It is noteworthy that the dispersivity of small-amplitude long waves tends to zero as $k \rightarrow 0$, since these waves are non-dispersive. In contrast, for short-wave packets, the dispersivity peaks at a wavenumber k_0 , where $|d_g|$ is maximal. Importantly, at $k_0 = 1$, the phase speed $c(k_0)$ remains $\mathcal{O}(1)$, which is comparable to the long-wave phase speed λ . Furthermore, since $k_0 |d_g| \ll |c|$, the group velocity c_g is close to c , and hence, also comparable to λ . Therefore, wave packets centred at $k_0 = 1$ travel at approximately the same speed as long waves, facilitating resonant coupling.

[Figure 15](#) shows space–time contours of \mathcal{S} and demonstrates it qualitatively corresponds to \mathcal{R} and TKE for all cases despite their fundamentally different mathematical definitions. For example, [figure 15\(a,b\)](#) suggests that in the SW and TW cases, \mathcal{S} is smallest in the central regions of the duct and, therefore, the relative velocity between packets of short waves and long waves is small in this region. Consequently, in SID, in the central region where long interfacial waves are in near resonance, they also resonate with energy-carrying short-wave packets, amplifying the growth of the unstable short waves, resulting

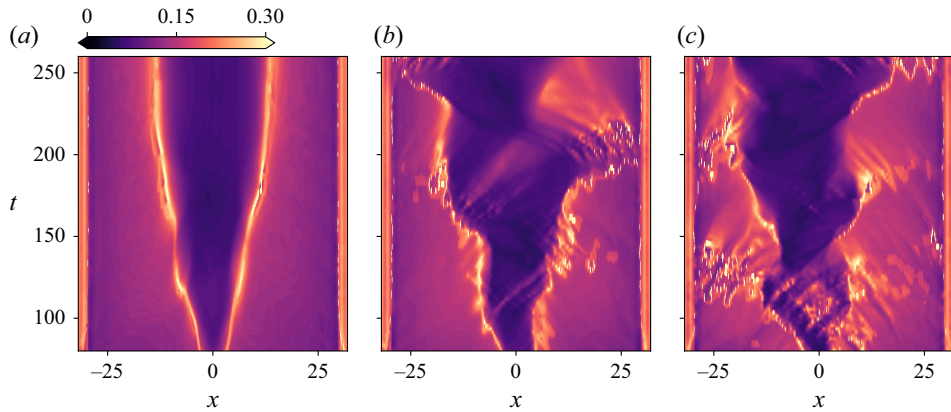


Figure 15. The resonance between long waves and short-wave packets measured by \mathcal{S} from (4.11): the spectral gap between most dispersive short-wave packets (i.e. wavenumbers centring around $k \approx 1$) and long waves in (a) SW, (b) TW and (c) I cases.

in breaking and generating small-scale turbulence and TKE. A more detailed comparison of the spatial distributions of TKE, \mathcal{R} and \mathcal{S} for the SW, TW and I cases at various times is presented in Appendix C.

Therefore, \mathcal{S} and \mathcal{R} serve as complementary diagnostics to unravel the route to turbulence and small-scale generation in a stratified exchange flow such as SID. The next section summarises these mechanisms.

4.5. Summary: resonance indicators and turbulence generation in space-time

In § 4.3 we introduced the indicator \mathcal{R} to quantify the long–long wave resonance, which we associate with large-scale energy transfer and the onset of instability in hydraulically controlled three-layer flows. This process is distinct from subcritical-to-supercritical hydraulic transitions, which do not require instability and are well known from single-layer theory. In § 4.4 we introduced \mathcal{S} to indicate small-scale generation through short-wave instability (long–short wave resonance). As shown in figures 12 and 15, both \mathcal{R} and \mathcal{S} have similar spatiotemporal distributions and both are consistent with the TKE distribution (see also Appendix C). Therefore, both long–long wave and long–short wave (packet) resonance contribute to the TKE and, consequently, influence the onset of turbulence in SID. The close similarity between the x – t distributions of \mathcal{R} and \mathcal{S} suggests an intrinsic connection between them, despite their fundamentally different mathematical definitions.

This process is also interpreted through the vertical structure of the wave modes. Long waves refer to disturbances with a small horizontal wavenumber k , corresponding to large-scale structures in the streamwise (x) direction. Because they vary slowly in x , the associated pressure and velocity perturbations extend over the full vertical depth of the duct, leading to broad vertical mode structures that align well with the background shear $U(z)$. This vertical coherence allows long waves to efficiently extract energy from the mean flow. However, their smooth structure induces weak vertical gradients, making them ineffective at generating localised shear or mixing across density interfaces.

Short waves, in contrast, are associated with a large horizontal wavenumber k and are confined near the density interfaces, with vertical structures that decay rapidly away from these regions. Although their projection onto the background shear is weaker, limiting their ability to extract energy directly from the mean flow, they can generate sharp interfacial

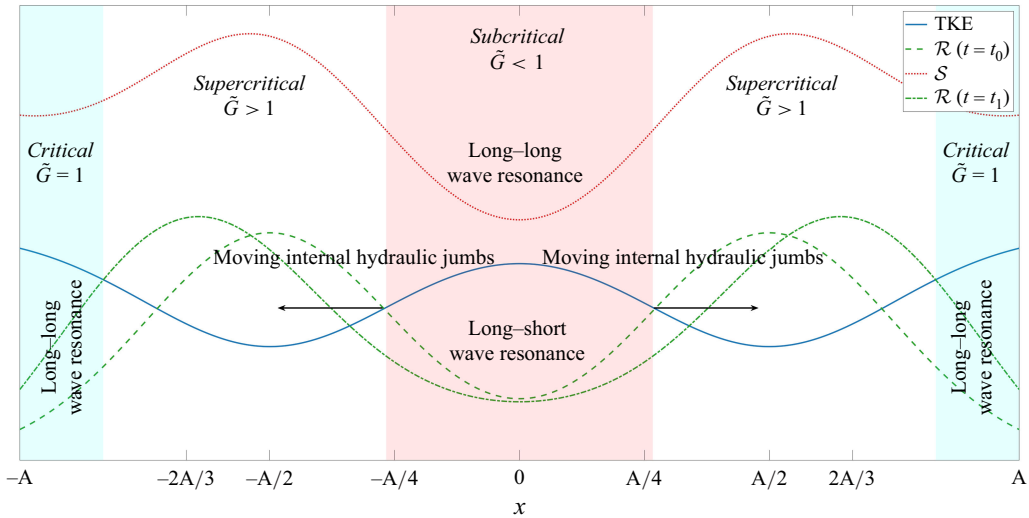


Figure 16. Conceptual profiles illustrating the instantaneous resonance indicators for long–long wave (\mathcal{R}) and long–short wave packet (S) interactions, shown alongside TKE for comparison. The parameter A represents the length-to-height aspect ratio of the stratified exchange flow considered here.

gradients that enhance local shear and promote mixing. Thus, they are more effective at driving the small-scale turbulent structures responsible for mixing across layers.

A key point revealed by our analysis is that long–short wave resonance enables a (non-linear) energy transfer from long waves to short waves. In regions where both \mathcal{R} and S are minimised, long waves that have extracted energy from the mean flow can resonantly transfer energy to short waves, thereby amplifying them and facilitating intense mixing. This interaction contributes to the observed generation of TKE and mixing in the middle layer.

This interpretation is supported by the qualitative similarity of low \mathcal{R} and S values with regions of enhanced TKE, as well as by previous theoretical work on long-wave-induced instability in exchange flows (Zhu *et al.* 2024). Together, these observations demonstrate the complementary roles of long and short waves in driving the transition to turbulence in stratified exchange flows such as SID.

The schematic profile in figure 16 illustrates the spatial distribution of long–long and long–short wave resonances within the SID flow, integrating the space–time interactions shown in figures 12 and 15 and comparing them with the TKE distribution presented in figure 9(a–c). This schematic also incorporates the spatial profiles in figure 20 and the analysis in Appendix C, providing insights into the identification of wave resonances and the hydraulic transition to turbulence. In the wave cases, as well as during the quiet phase of the I case, critical regions near the duct ends (where $\tilde{G} \approx 1$) are observed, where the lowest values of \mathcal{R} coincide with the strongest TKE, while S remains minimised. This pattern suggests that hydraulic transition and TKE generation are primarily driven by long–long wave resonance, in conjunction with short-wave instabilities at the upper and lower interfaces without substantial resonant interactions. In the duct centre (subcritical region where $\tilde{G} < 1$), both long–long and long–short wave resonances coexist, facilitating the hydraulic transition to turbulence. As the central hydraulic jumps progress toward the duct ends and the subcritical region expands, these regions of resonant interaction grow, occupying an increasingly extensive portion of the duct.

4.6. Application to oceanic flows

The theoretical framework developed in this study, based on three-layer hydraulic analysis and the diagnosing resonant interactions between long and short waves, can be applied to stratified exchange flows in the ocean. These include confined regions like straits, estuaries, fjords, canyons and semi-enclosed basins, where salinity and/or temperature stratification often create three-layer flows governed by hydraulic controls. A notable example is the Bosphorus Strait, linking the Black Sea and the Mediterranean, where bathymetric constriction and density differences support layered, hydraulically controlled stratified exchange. Below, we outline several potential applications.

4.6.1. Implementations for adaptive resolution in ocean models

Recent advances in high-resolution numerical ocean modelling, particularly the lat-lon-cap (LLC) simulations such as LLC4320 (Gallmeier *et al.* 2023), have enabled global ocean simulations at approximately $1/48^\circ$ horizontal resolution (~ 2 km at the equator) using the Massachusetts Institute of Technology general circulation model (MITgcm). Despite this relatively high spatial resolution, turbulence is far from being resolved, and mixing remains parameterised through local gradient-based closures.

The diagnostic measures \mathcal{R} and \mathcal{S} developed in this study quantify the conditions under which long waves resonate with dispersive short-wave packets. Since these diagnostics rely only on layer-averaged quantities – velocities, densities and interface displacements – they are directly applicable to model outputs.

In coastal and marginal seas, particularly in estuarine and strait-like geometries where the Rossby number is large and rotational effects are weak, the three-layer approximation proposed in this study is especially appropriate. Our results imply that even without directly resolving turbulence, ocean models can leverage these diagnostics to identify zones of intensified TKE generation. This insight may be valuable for implementing adaptive mesh refinement in regional models, where local grid refinement could be concentrated in dynamically active zones suggested by long–short wave resonance conditions. Because all diagnostic expressions derived here are analytical, this integration poses no additional computational burden and presents a promising route toward physics-informed resolution enhancement in estuarine and strait flows.

4.6.2. Implications for mixing parametrisation

While in this study we applied the theoretical framework to DNS datasets of SID flows, the methodology can be extended to ocean circulation models, which typically solve shallow water or layer-averaged primitive equations. These models do not resolve turbulence explicitly and usually assume hydrostatic balance, thus neglecting non-hydrostatic pressure gradients that play a key role in wave-driven turbulence and mixing.

The model output can be decomposed into three effective layers by discretising the vertical structure into bulk velocity, density and thickness in each layer via appropriate integration. These layer-averaged quantities serve as input to the theory developed here, enabling the computation of resonance diagnostics \mathcal{R} and \mathcal{S} , the linear wave speed c , the imaginary part c_i as a proxy for instability and the dispersivity $|d_g|$, defined as the magnitude of the group velocity derivative with respect to wavenumber. In regions where \mathcal{S} is locally minimum or $|d_g|$ is large, we suggest that turbulent viscosity in mixing parametrisations could adopt a velocity scale based on the group velocity of the most dispersive wave packet, $c_g(k_0)$, where k_0 corresponds to the wavenumber at which $|d_g|$ peaks.

This provides a leading-order correction to existing mixing schemes that rely solely on local shear or stratification. It leverages the resolved mean state of an ocean model

to estimate non-hydrostatic, wave-induced pathways to turbulence, which are otherwise absent in hydrostatic model formulations. Consequently, it offers a route to enhance parametrisation schemes with minimal computational overhead while capturing key non-local mixing mechanisms.

4.7. Outlook

The good agreement between DNS and temperature-stratified shadowgraph experiments in Zhu *et al.* (2023) suggests that hydraulic effects, including both long–long and long–short wave interactions, play an important role in these experiments. Furthermore, recent augmented experimental data (Zhu *et al.* 2024) from salt-stratified experiments using physics-informed neural networks and the reconstructed pressure field indicates that the offset of the density interface, and thus, the asymmetry in short (Holmboe) waves are also caused by hydraulic effects. However, the question of how the Prandtl number influences these interactions remains open.

A promising direction for future work is the inclusion of a steady or oscillatory barotropic background flow. In many geophysical contexts, particularly in estuaries and straits, the exchange flow is influenced by persistent barotropic currents or low-frequency tidal forcing. A key assumption in the present formulation is that the net volume flux vanishes, i.e. $u_1 h_1 + u_2 h_2 + u_0 h_0 = 0$. In the presence of barotropic forcing, this condition generalises to $u_1 h_1 + u_2 h_2 + u_0 h_0 = Q(x, t)$, where Q is a prescribed barotropic volume flux. This introduces spatiotemporal variability into the base state and alters the characteristics of the three-layer flow. Consequently, resonance conditions and growth rates become time dependent on the scale relevant to the forcing time scale, potentially modifying the nature and location of mixing.

Another natural extension is to incorporate rotation. While the present study assumes non-rotating flows – a valid assumption in narrow, laterally confined geometries – large-scale oceanic flows are influenced by Coriolis effects, including continental shelf regions, curved straits and deep ocean canyons. Including rotation would introduce geostrophic balance into the mean flow and modify the structure of internal wave characteristics. This could suppress or enhance wave interactions depending on the local Rossby number and geometry.

These extensions, accounting for barotropic variability and rotational effects, represent important steps toward adapting the resonance-based framework developed here into a more general predictive tool for mixing in real oceanic flows.

5. Conclusions

This paper has discussed the three-layer hydraulics of stratified exchange flows in inclined ducts and significantly improved upon the two-layer hydraulics of Atoufi *et al.* (2023). The improvement lies in the inclusion of viscous effects, non-hydrostatic pressure gradients and a third mixed layer in the quantitative analysis of the internal hydraulics and wave instabilities, which is generalisable to any three-layer exchange flows with mixing.

We first introduced the viscous three-layer-averaged equations with non-hydrostatic pressure gradients and mixing. We then provided expressions for the eigenvalues of the homogeneous system of equations (i.e. characteristics) and the propagation of information, and defined a composite Froude number based on the characteristics of the three-layer flows. While related formulations have appeared in prior work, this approach based on characteristics has not, to our knowledge, been previously reported. Unlike two-layer hydraulics, three-layer hydraulics predict a subcritical/critical region, separated by two hydraulic control points, within which flow transitions from supercritical to

subcritical, and is accompanied by TKE generation. We have shown theoretically that the imaginary component of characteristics represents the growth rate of the linear long waves propagating on a background three-layer flow.

The system was then decomposed into three isolated layers through matrix decomposition, where the characteristics are strictly real. We demonstrated that long-wave growth occurs when interactions between these isolated layers are allowed. From this decomposition, perturbed upper and lower layers were defined resulting from the interaction between the isolated upper and lower layers, and influenced by the isolated middle layer. The spatiotemporal TKE was linked to long-wave resonance by quantifying the spectral gaps between the characteristics of the perturbed upper and lower layers. Additionally, we connected TKE to the resonance of dispersive short-wave packets carrying energy, characterised by a wavelength three times the duct height, whose group velocity drives kinetic energy transfer.

The first turbulent transition mechanism arises when long waves have identical speeds and growth rates, facilitating resonance between them. In this case, the spectral gap between the perturbed upper and lower layers vanishes. Such a resonant interaction also changes the base state around which the perturbation evolves and triggers instabilities. The next mechanism arises when long waves directly interact with a packet of dispersive short waves, where the phase speed and growth rate of their amplitude envelope are equal to those of long interfacial waves. In this case, direct interactions between long and short waves trigger instability and the generation of smaller scales.

This paper has focused on the three regimes identified in Zhu *et al.* (2023): stationary and travelling wave regimes (SW and TW case) together with the intermittently turbulent regime (I case). We provided the first evidence from three-dimensional DNS datasets that strong TKE correlates with long–long waves and long–short wave packet resonance. We deduce that the transition to turbulence in stratified exchange flow exhibits a high degree of non-locality owing to resonant interactions between long and short waves. The parametrisation of mixing in applications, including in the ocean, is often based on local gradients of the flow. This paper underscores the need to revisit modelling strategies to account for non-local pathways to turbulence, especially in situations where the flow is hydraulically controlled or influenced, and susceptible to long-wave instability (e.g. in estuaries and straits).

Funding. We acknowledge support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation Grant No 742480 ‘Stratified Turbulence And Mixing Processes’ (STAMP). Parts of the simulations were carried out with resources from Compute/Calcul Canada. A.L. is funded by a NERC Independent Research Fellowship (NE/W008971/1). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Declaration of interests. The authors report no conflict of interest.

Appendix A. Linear instabilities and the link to three-layer hydraulics

We assume three-layer piecewise constant base velocity and density profiles of the form

$$\mathcal{U}(z) = \begin{cases} u_1 & \frac{h_0}{2} < z \leq h_1 + \frac{h_0}{2}, \\ u_0 & -\frac{h_0}{2} < z < \frac{h_0}{2}, \\ u_2 & -h_2 - \frac{h_0}{2} \leq z < -\frac{h_0}{2}, \end{cases} \quad \mathcal{R}(z) = \begin{cases} \rho_1 & \frac{h_0}{2} < z \leq h_1 + \frac{h_0}{2}, \\ \rho_0 & -\frac{h_0}{2} < z < \frac{h_0}{2}, \\ \rho_2 & -h_2 - \frac{h_0}{2} \leq z < -\frac{h_0}{2}, \end{cases} \quad (\text{A1})$$

where we use layer-average quantities from DNS data to compute $\mathcal{U}(z)$ and $\mathcal{R}(z)$. We add a perturbation streamfunction of the form $\widehat{\psi}(z)\exp ik(x - ct)$ to represent linear disturbance waves propagating in a background flow with prescribed $\mathcal{U}(z)$ and $\mathcal{R}(z)$. We implicitly assume that the above parallel base flow is a good model for exchange flows varying slowly along x compared with the wavelength of the perturbation. The evolution of the perturbation inside the duct (away from the inlet/outlet) is given by the forced Taylor–Goldstein equation (TGE) (Atoufi *et al.* 2023):

$$[(\mathcal{U} - c)\widehat{\psi}' - \mathcal{U}'\widehat{\psi}]' - (\mathcal{U} - c)k^2\widehat{\psi} - \frac{Ri \cos \theta \mathcal{R}'}{\mathcal{U} - c}\widehat{\psi} = -\frac{i}{k}Ri \sin \theta \left[\frac{\mathcal{R}'\widehat{\psi}}{\mathcal{U} - c} \right]', \quad (\text{A2})$$

where primes denote d/dz . For the three-layer base flow (A1), $\mathcal{U}'' = \mathcal{R}' = 0$ everywhere except at the interface, the TGE (A2) reduces to

$$\widehat{\psi}'' - k^2\widehat{\psi} = 0 \quad (\text{A3})$$

everywhere except at the lower and upper interfaces and we therefore seek the solution of the form

$$\widehat{\psi} = \begin{cases} M_1 \sinh k \left(h_1 + \frac{h_0}{2} - z \right) & \frac{h_0}{2} < z \leq h_1 + \frac{h_0}{2}, \\ M_2 \sinh k(z) + M_3 \cosh k(z) & -\frac{h_0}{2} < z < \frac{h_0}{2}, \\ M_4 \sinh k \left(h_2 + \frac{h_0}{2} + z \right) & -h_2 - \frac{h_0}{2} \leq z < -\frac{h_0}{2}. \end{cases} \quad (\text{A4})$$

We note this solution satisfies the impermeability condition at $z = -h_2 - (h_0/2)$ and $z = h_1 + (h_0/2)$ where $\widehat{w} = ik\widehat{\psi} = 0$ since the vertical velocity of the disturbance waves vanishes at the upper and lower boundaries. We also note that

$$\begin{aligned} \mathcal{U}' &= (u_1 - u_0) \delta \left(z - \frac{h_0}{2} \right) + (u_0 - u_2) \delta \left(z + \frac{h_0}{2} \right), \\ \mathcal{R}' &= (\rho_1 - \rho_0) \delta \left(z - \frac{h_0}{2} \right) + (\rho_0 - \rho_2) \delta \left(z + \frac{h_0}{2} \right), \end{aligned} \quad (\text{A5})$$

where δ is the Dirac delta function and $\mathcal{U}'' = \mathcal{R}' = 0$ everywhere except at the lower and upper interfaces. The matching conditions are derived by integrating (A2) over the neighbourhood of the lower and upper interfaces ($z = \pm h_0/2$) leading to

$$\begin{aligned} \left[\left[\frac{\widehat{\psi}}{\mathcal{U} - c} \right] \right]_{\eta_i} &= 0, \\ \left[[(\mathcal{U} - c)\widehat{\psi}'] \right]_{\eta_i} + g_i \left(\frac{\widehat{\psi}}{\mathcal{U} - c} \right)_{z=\eta_i} &= 0, \end{aligned} \quad (\text{A6})$$

where η_i ($i = 1, 2$) refers to lower and upper interface locations as determined in § 2.2. In deriving the matching conditions (A6) we also considered the small angle limit where $\sin(\theta) \ll \cos \theta$ and neglected the right-hand side of (A2). The first and second conditions guarantee the continuity of normal velocity and pressure fluctuations, respectively, across the interfaces. Equation (A6) provides a 4×4 system of algebraic equations that we solve to find M_1, M_2, M_3 , and M_4 . The condition for the non-trivial solution of (A6) leads to the phase velocity of the linear disturbance waves to satisfy the general quartic equation $\sum_{n=0}^4 b_n(\mathbf{q}, k)c^n = 0$, where the coefficients are introduced in the next section.

A.1. Coefficients in quartic equation derived from three-layer TGE

To manage the complexity of the coefficients b_n in the quartic dispersion relation, we introduce hierarchical auxiliary variables grouped in three decks, facilitating compact representation and analytical traceability.

The upper deck contains the main coefficients b_n appearing directly in the dispersion relation

$$\begin{aligned}
 b_4 &= \sigma_3, \\
 b_3 &= \sinh(k(h_0 - h_1 + h_2))(u_0 - u_1) + \sinh(k(h_0 + h_1 - h_2))(u_0 - u_2) \\
 &\quad - \sigma_3(2u_0 + u_1 + u_2), \\
 b_2 &= u_0^2(\alpha_1 + \alpha_2 + 6\alpha_4) + u_1^2(\alpha_1 + \alpha_3) + u_2^2(\alpha_2 + \alpha_3) + 4u_0(u_1\alpha_1 + u_2\alpha_2) \\
 &\quad + 4u_1u_2\alpha_3 + \gamma_1 + \gamma_2, \\
 b_1 &= (1 - 2\sigma_1^2)(\gamma_7 + \gamma_9) - \sigma_1\sigma_2\gamma_8, \\
 b_0 &= u_0^2\gamma_3 + 2u_0^4\sigma_1\beta_1\sigma_2 + 2u_1^2u_2^2\beta_4\sigma_1\sigma_2 + u_2^2\gamma_4 + u_1^2\gamma_5 + \gamma_6.
 \end{aligned} \tag{A7}$$

The middle deck collects all hyperbolic expressions that depend on the layer depths and wavenumber. These consist of combinations of sinh and cosh terms and provide the building blocks of the system

$$\begin{aligned}
 \sigma_1 &= \cosh\left(\frac{h_0k}{2}\right), \quad \sigma_2 = \sinh\left(\frac{h_0k}{2}\right), \quad \sigma_3 = \sinh(k(h_0 + h_1 + h_2)), \\
 \alpha_1 &= \cosh(h_0k) \cosh(h_1k) \sinh(h_2k), \quad \alpha_2 = \cosh(h_0k) \cosh(h_2k) \sinh(h_1k), \\
 \alpha_3 &= \cosh(h_1k) \cosh(h_2k) \sinh(h_0k), \quad \alpha_4 = \sinh(h_0k) \sinh(h_1k) \sinh(h_2k), \\
 \alpha_5 &= \cosh(h_0k) \sinh(h_1k) \sinh(h_2k), \quad \alpha_6 = \cosh(h_2k) \sinh(h_0k) \sinh(h_1k), \\
 \alpha_7 &= \cosh(h_1k) \sinh(h_0k) \sinh(h_2k), \\
 \beta_1 &= \sinh(h_1k) \sinh(h_2k), \quad \beta_2 = \cosh(h_1k) \sinh(h_2k), \\
 \beta_3 &= \cosh(h_2k) \sinh(h_1k), \quad \beta_4 = \cosh(h_1k) \cosh(h_2k).
 \end{aligned} \tag{A8}$$

The lower deck expresses all composite terms that couple the middle-deck hyperbolic functions with velocities and gravitational accelerations. This layer systematically simplifies and organises the structure of b_n

$$\begin{aligned}
 \gamma_1 &= -\frac{g_1(\alpha_5 + \alpha_6)}{k}, \quad \gamma_2 = -\frac{g_2(\alpha_5 + \alpha_7)}{k}, \quad \gamma_3 = \frac{(1 - 2\sigma_1^2)\beta_1(g_1 + g_2)}{k}, \\
 \gamma_4 &= -\frac{2g_1\beta_3\sigma_1\sigma_2}{k}, \quad \gamma_5 = -\frac{2g_2\beta_2\sigma_1\sigma_2}{k}, \quad \gamma_6 = \frac{2g_1g_2\sigma_1\beta_1\sigma_2}{k^2}, \\
 \gamma_7 &= 2(u_0u_1^2 + u_0^2u_1)\beta_2 + 2(u_0u_2^2 + u_0^2u_2)\beta_3 + \frac{2u_0\beta_1(g_1 + g_2)}{k}, \\
 \gamma_8 &= 8u_0^3\beta_1 + 4u_1u_2(u_1 + u_2)\beta_4 - u_2\frac{4g_1\beta_3}{k} - u_1\frac{4g_2\beta_2}{k}, \\
 \gamma_9 &= (1 - 2\sigma_1^2)(u_1^2\beta_2 - u_2^2\beta_3).
 \end{aligned} \tag{A9}$$

A.2. Dispersion relation

The dispersion relation for the small tilt angle such that $\sin \theta \ll \cos \theta$ can be expressed in the form

$$\sum_{n=0}^4 b_n(\mathbf{q}, k)c^n = 0, \tag{A10}$$

where the wavenumber-dependent coefficients b_n are given in Appendix A.1 and the analytical solution to (A10) is obtained by replacing a_n with b_n in (3.9) whose solution is given in (3.12). Note that the two solutions $c_{1,2}$ (the first and second root of (A10)) are the speed of linear waves on the upper interface while the two solutions $c_{3,4}$ (the third and fourth root of (A10)) are those on the lower interface. In the long-wave limit where $kh_1, kh_0, kh_2 \rightarrow 0$, we find that $b_i \rightarrow ka_i$ ($i = 0, \dots, 3$) and that

$$c_i(k) \longrightarrow \lambda_i \quad \forall i \in \{1, 2, 3, 4\} \quad \text{when } k \ll 1. \tag{A11}$$

Therefore, c_1, c_2, c_3 and c_4 individually converge to $\lambda_1, \lambda_2, \lambda_3$ and λ_4 in the long-wave limit. The characteristics are thus the phase speed of the linear long waves of small amplitude disturbing individual interfaces. Consequently, the imaginary part of the characteristics λ^I of the three-layer flow also describes the potential growth ($\lambda^I > 0$), decay ($\lambda^I < 0$) or neutral state ($\lambda^I = 0$) of these linear long waves on each interface.

Although we neglected horizontal buoyancy forcing in the three-layer dispersion relation, it is noteworthy to comment that in the proximity of the critical layer this forcing dominates and shapes the structure of the unstable modes. This influence is attributed to the vertical derivative present on the right-hand side of (A2), leading to the appearance of $(U - c)^{-2}$ terms, particularly evident in the long-wave limits (owing to the $1/k$ prefactor). In future work we will further investigate the properties of the solution to (A2) for various wavenumbers near the critical layers.

A.3. Pure exchange flow with stagnant middle layer

In the limit $u_0 = 0, u_2 = -u_1, h_2 = h_1$ and $g_2 = g_1$, corresponding to a symmetric exchange flow with a stagnant middle layer, the dispersion relation coefficients simplify to

$$\begin{aligned} b_4 &= \sinh(k(h_0 + 2h_1)), \quad b_3 = 0, \quad b_1 = 0, \\ b_2 &= \frac{2}{k} [g_1 \sinh(h_1 k) \sinh(k(h_0 + h_1)) - ku_1^2 \sinh(k(h_0 - h_1)) \cosh(h_1 k)], \\ b_0 &= \frac{2}{k} [g_1 u_1^2 \sinh(2h_1 k) + k(g_1^2 \sinh^2(h_1 k) + u_1^4 \cosh^2(h_1 k))] \cosh\left(\frac{h_0 k}{2}\right) \sinh\left(\frac{h_0 k}{2}\right). \end{aligned} \tag{A12}$$

For long waves ($k \ll 1$), the phase speed converges to the analytical expression given in (3.15), becoming effectively independent of k and, hence, non-dispersive. In this regime, the numerically computed branches λ_1^I and λ_4^I , as illustrated in figure 5, are indistinguishable from long-wave speeds c_1^I and c_4^I , where $k = 0.001$ is chosen to represent the long-wave limit.

A.4. Application to the data

The growth rate of the linear waves on the upper interface c_2^I and the growth rate of the linear waves on the lower interface c_4^I are shown in figure 17(a,b) for the TW case (left column) and the I case (right column) at a representative time of $t = 180$. The fastest growth in the upper (lower) interfacial wave in the TW case occurs in the region where $x \leq -15$ ($x \geq 15$). In the I case, although both upper and lower interfacial waves are unstable almost everywhere inside the duct, the $x \geq 15$ and $x \leq -15$ regions exhibit the fastest growth rates. In both cases, both long waves ($k \ll 1$) and short waves ($k \gg 1$) are unstable. The range of wavenumbers with the fastest growth rate is $0.1 \leq k \leq 2$, with the peak around $k \approx 1$. Consistently, many previous linear studies showed saturation in the growth rate of short waves for $k = 1.5$, leading to Kelvin–Helmholtz and Holmboe

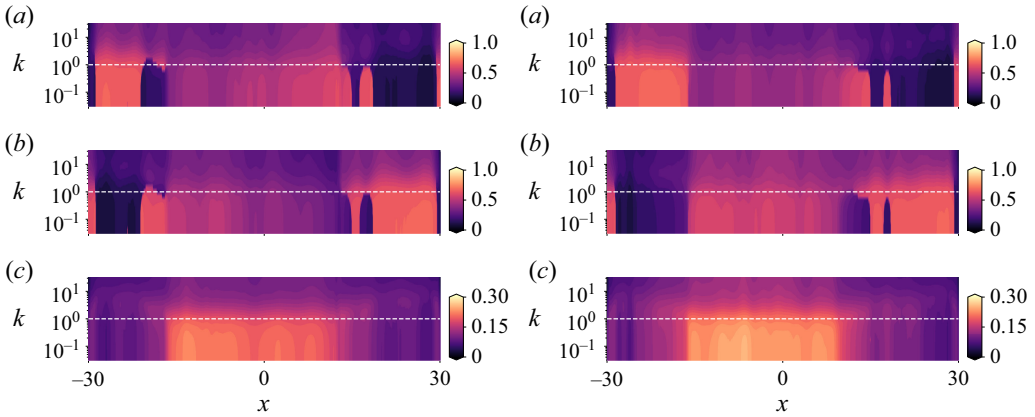


Figure 17. Dispersion relation of the linear interfacial waves for the TW case (left column) and the I case (right column) at $t = 180$. Based on the analytical solution given in (A10) and (A9), we show (a) the lower interface growth rate c_2^I , (b) the upper interface growth rate c_4^I and (c) the amplification of the growth rate $c_2^I c_4^I$.

instabilities, depending on the thickness of the density interface (Smyth & Peltier 1991; Lefaube *et al.* 2018; Ducimetière *et al.* 2021).

In figure 17(c) the product $c_2^I c_4^I$ is shown to highlight the regions in the flow with simultaneous amplification of the growth rates of the interfacial waves. The product $c_2^I c_4^I$ is largest for medium and long waves ($k < 1$) in all cases. In the I case, this simultaneous amplification occurs over a more extended region inside the duct compared with the TW case, where it is more localised near the centre at $x = 0$, consistent with the large values of TKE in figure 9 at the duct centre.

The dispersivity d_g of the I case at various x locations are shown in figure 18(a,b) (active stage) and figure 18(c,d) (quiet stage). The d_g plot for the SW and TW cases is overall similar to the quiet stage of the I case and is not shown here for brevity. We observe that $d_g \rightarrow 0$ in the long-wave limit ($k \rightarrow 0$) as expected. The most dispersive waves are short, with $1 \leq k \leq 2$ for all x and t . This range of wavenumbers also corresponds to the fastest growth in the amplitude of the linear waves, as shown in the right columns of figure 17(a,b) for these short waves. In the quiet stage, the dispersivity of the waves is almost independent of x away from the inlet and outlet (figure 18c,d). In the active stage, however, the spatial variation in d_g is more pronounced.

Appendix B. Spectral gap and long–long-wave resonance

The concept of the spectral gap is well established in matrix perturbation theory (see, for example, Stewart 1975; Elsner & Sun 1982). In this section we review this concept for completeness and apply it within the context of a three-layer system and the resonance between long waves in the subsystems defined in the building-block (matrix) decomposition of (4.1).

B.1. Three-layer system decomposition and the spectral gap

We consider a real matrix pair, denoted \mathbf{M} , with dimensions 6×12 , associated with a three-layer system. Using figure 19, we represent its characteristics, $\lambda(\mathbf{M}) \in \mathbb{C}$, which consist of real and imaginary components, as

$$\lambda(\mathbf{M}) = \lambda^R(\mathbf{M}) + i\lambda^I(\mathbf{M}) = (\lambda^R(\mathbf{M}), \lambda^I(\mathbf{M}), 0) = \mathbf{P}_c. \quad (\text{B1})$$

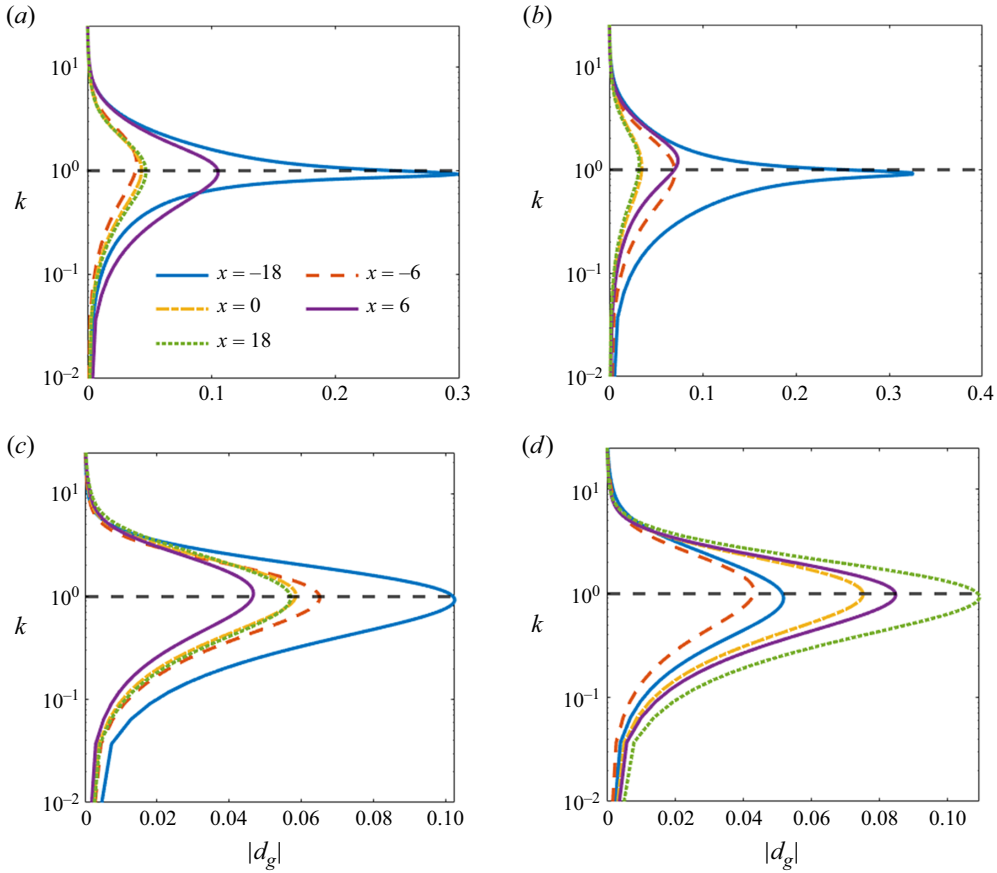


Figure 18. Dispersivity values $|d_g|$ of disturbance waves in case I at various x locations at $t = 110$ during the active stage (a,b) and at $t = 180$ in the quiet stage (c,d). In panels (a,c) the $|d_g|$ of the upper interfacial waves is presented, while panels (c,d) show the dispersivity of the lower interfacial waves.

Next, we apply stereographic projection to map these characteristics onto the unit sphere. Stereographic projection is a standard technique for mapping points from a plane – in this case, the complex plane of eigenvalues – onto a (Riemann) sphere. This projection enables a geometrically meaningful analysis of the matrix’s spectral properties, as each point on the sphere corresponds to a pair of real and imaginary components from the complex plane, revealing insights into the eigenvalue structure. This approach is also geometrically instructive for examining $\lambda(\mathbf{M})$ against a second matrix pair with strictly real characteristics, which vary within the range $[-1, 1]$, as is the case, for example, for $\lambda(\mathbf{Z})$.

We define a point on the sphere as $\mathbf{P} = (\lambda_1^R, \lambda_1^I, \lambda(\mathbf{Z}_1))$ that is the stereographic projection of \mathbf{P}_c , where the unit sphere is constructed such that

$$(\lambda_1^R)^2 + (\lambda_1^I)^2 + \lambda(\mathbf{Z}_1)^2 = 1. \tag{B2}$$

The north pole of the sphere is denoted by $\mathbf{N} = (0, 0, 1)$, and the line connecting \mathbf{N} to the point $\lambda(\mathbf{M})$ on the complex plane is given by the parametric equation

$$\begin{aligned}
 (\lambda_1^R, \lambda_1^I, \lambda(\mathbf{Z}_1)) &= (1-s)\lambda(\mathbf{M}) + sN \\
 &= ((1-s)\lambda^R(\mathbf{M}), (1-s)\lambda^I(\mathbf{M}), s), \quad \forall s \in \mathbb{R} - \{1\}. \quad (\text{B3})
 \end{aligned}$$

Here, s is a parameter that controls the position along the line between the point $\lambda(\mathbf{M})$ and the north pole N . When $s = 0$, the point is at $\lambda(\mathbf{M})$, and when s approaches unity the point reaches the north pole N . If one takes the magnitude of both sides of (B3) and substitutes the condition of the unit sphere, given by (B2), we relate the square of the distance between the projection point and the origin (i.e. the squared magnitude of $\lambda(\mathbf{M})$ in the complex plane) to the parameter s :

$$|\lambda(\mathbf{M})|^2 = (\lambda^R(\mathbf{M}))^2 + (\lambda^I(\mathbf{M}))^2 = \frac{1+s}{1-s}. \quad (\text{B4})$$

This result shows that the stereographic projection maps points from the complex plane onto the sphere in such a way that the distance from the origin in the complex plane, $|\lambda(\mathbf{M})|$, is related to the projection parameter s by the formula $(1+s)/(1-s)$. This relationship provides a geometric interpretation of how the spectral characteristics of the matrix are represented on the sphere. To do so, we solve for s to obtain

$$s = \frac{|\lambda(\mathbf{M})|^2 - 1}{|\lambda(\mathbf{M})|^2 + 1}. \quad (\text{B5})$$

We now substitute (B5) into (B3) to express the real and imaginary components of the projection on the sphere:

$$\lambda_1^R = \frac{2\lambda^R(\mathbf{M})}{1 + |\lambda(\mathbf{M})|^2}, \quad \lambda_1^I = \frac{2\lambda^I(\mathbf{M})}{1 + |\lambda(\mathbf{M})|^2}, \quad \lambda(\mathbf{Z}_1) = \frac{|\lambda(\mathbf{M})|^2 - 1}{|\lambda(\mathbf{M})|^2 + 1}. \quad (\text{B6})$$

The stereographic projection allows us to map all points $(\lambda_1^R, \lambda_1^I, \lambda(\mathbf{Z}))$ on the sphere to corresponding points in the complex plane \mathbb{C} , constructed from the eigenvalues of \mathbf{M} .

Next, consider another point $\mathbf{Q} = (\lambda_2^R, \lambda_2^I, \lambda(\mathbf{Z}_2))$ on the sphere, where

$$(\lambda_2^R)^2 + (\lambda_2^I)^2 + \lambda(\mathbf{Z}_2)^2 = 1, \quad (\text{B7})$$

which connects N to $\lambda(\mathbf{M} + \mathbf{M}')$. We can write a similar expression to relate the components of \mathbf{Q} to $\lambda^R(\mathbf{M} + \mathbf{M}')$ and $\lambda^I(\mathbf{M} + \mathbf{M}')$ as

$$\lambda_2^R = \frac{2\lambda^R(\mathbf{M} + \mathbf{M}')}{1 + |\lambda(\mathbf{M} + \mathbf{M}')|^2}, \quad \lambda_2^I = \frac{2\lambda^I(\mathbf{M} + \mathbf{M}')}{1 + |\lambda(\mathbf{M} + \mathbf{M}')|^2}, \quad \lambda(\mathbf{Z}_2) = \frac{|\lambda(\mathbf{M} + \mathbf{M}')|^2 - 1}{|\lambda(\mathbf{M} + \mathbf{M}')|^2 + 1}. \quad (\text{B8})$$

The Euclidean distance $|\mathbf{PQ}|$ between \mathbf{P} and \mathbf{Q} is

$$\begin{aligned}
 |\mathbf{PQ}| &= \sqrt{(\lambda_1^R - \lambda_2^R)^2 + (\lambda_1^I - \lambda_2^I)^2 + (\lambda(\mathbf{Z}_1) - \lambda(\mathbf{Z}_2))^2} \\
 &= \sqrt{2 - 2(\lambda_1^R \lambda_2^R + \lambda_1^I \lambda_2^I + \lambda(\mathbf{Z}_1)\lambda(\mathbf{Z}_2))} \quad (\text{B9})
 \end{aligned}$$

due to the constraints (B2) and (B7), noting that the term inside the bracket yields $\mathbf{P} \cdot \mathbf{Q}$ that is equal to the cosine of the angle between the two vectors connecting \mathbf{P} and \mathbf{Q} to the origin. Using the expressions from (B6) and (B8), we conclude that

$$\lambda_1^R \lambda_2^R + \lambda_1^I \lambda_2^I + \lambda(\mathbf{Z}_1)\lambda(\mathbf{Z}_2) = 1 - \frac{2|\lambda(\mathbf{M} + \mathbf{M}') - \lambda(\mathbf{M})|^2}{(|\lambda(\mathbf{M} + \mathbf{M}')|^2 + 1)(|\lambda(\mathbf{M})|^2 + 1)}, \quad (\text{B10})$$

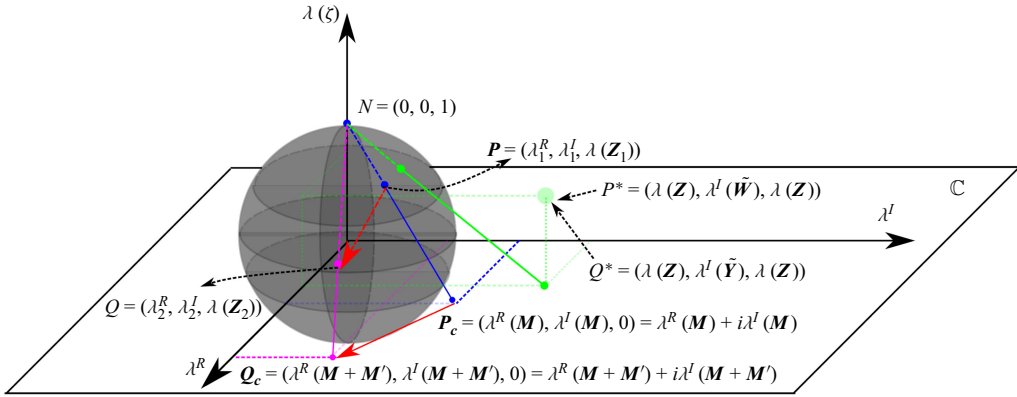


Figure 19. Spectral variation between two matrix pairs, \mathbf{M} and $\mathbf{M} + \mathbf{M}'$, measured by \mathbf{QP} , the back projection of $\mathbf{Q}_c \mathbf{P}_c$ onto the unit sphere that encompasses the spectrum of a third matrix pair, ζ , whose eigenvalues satisfy $\lambda(\zeta) \in [-1, 1]$. The north pole on the unit sphere is denoted by $N = (0, 0, 1)$, and $\mathbf{Q}_c \mathbf{P}_c$ represents the stereographic projection of \mathbf{QP} . The spectral gap, $S_{\mathbf{M}}(\mathbf{M}') = |\mathbf{QP}|/2$, quantifies the magnitude of spectral variation between \mathbf{M} and $\mathbf{M} + \mathbf{M}'$. The resonance condition requires $P^* = Q^*$.

which leads to the final form of the Euclidean distance:

$$|\mathbf{PQ}| = \frac{|\lambda(\mathbf{M} + \mathbf{M}') - \lambda(\mathbf{M})|}{\sqrt{|\lambda(\mathbf{M} + \mathbf{M}')|^2 + 1} \sqrt{|\lambda(\mathbf{M})|^2 + 1}}. \tag{B11}$$

This expression represents the distance between two arbitrary points on the Riemann sphere based on the spectra of two arbitrary matrices. The matrices \mathbf{M} and $\mathbf{M} + \mathbf{M}'$ are square matrix pairs of size 6×12 , each having six eigenvalues. Based on (B11), we can construct the spectral variation matrix as

$$d_{jk}(\mathbf{M} + \mathbf{M}', \mathbf{M}) = \frac{\lambda_j(\mathbf{M} + \mathbf{M}') - \lambda_k(\mathbf{M})}{\sqrt{|\lambda_j(\mathbf{M} + \mathbf{M}')|^2 + 1} \sqrt{|\lambda_k(\mathbf{M})|^2 + 1}}, \quad 1 \leq j, k \leq 4, \tag{B12}$$

which measures variation among all characteristics of \mathbf{M} and $\mathbf{M} + \mathbf{M}'$. The spectral variation between \mathbf{M} and $\mathbf{M} + \mathbf{M}'$ is also illustrated in figure 19 by the magnitude of the the vector \mathbf{QP} as the back projection of $\mathbf{Q}_c \mathbf{P}_c$ onto the unit sphere, encompassing $-1 \leq \lambda(\zeta) \leq 1$. For each row in d_{jk} , we search for the k that results in the minimum column-wise values for $\mathcal{D}_j = \min \{|d_{j1}|, |d_{j2}|, |d_{j3}|, |d_{j4}|\}$ to construct a vector from the columns of d_{jk} that yields the least spectral variation in the j th row. The spectral gap is then the infinity norm $\|\mathcal{D}\|_{\infty}$. Collectively, the spectral gap is computed as

$$S_{\mathbf{M}}(\mathbf{M}') \equiv \max_j \min_k |d_{jk}(\mathbf{M} + \mathbf{M}', \mathbf{M})|, \quad 1 \leq j, k \leq 4, \tag{B13}$$

which is consistent with the definition of Stewart (1975) and Elsner & Sun (1982).

B.2. Long-wave resonance indicator

Let us now substitute the pairs \mathbf{M} and $\mathbf{M} + \mathbf{M}'$ with \mathbf{W} and \mathbf{Z} , representing the isolated upper and middle layer system defined in (4.5). Using (B13), the quantity $S_{\mathbf{Z}}(\mathbf{W})$ measures the spectral gap between \mathbf{W} and \mathbf{Z} . Similarly, $S_{\mathbf{Z}}(\mathbf{Y})$ measures the spectral gap between \mathbf{Y} the isolated lower layer and \mathbf{Z} the isolated middle layer.

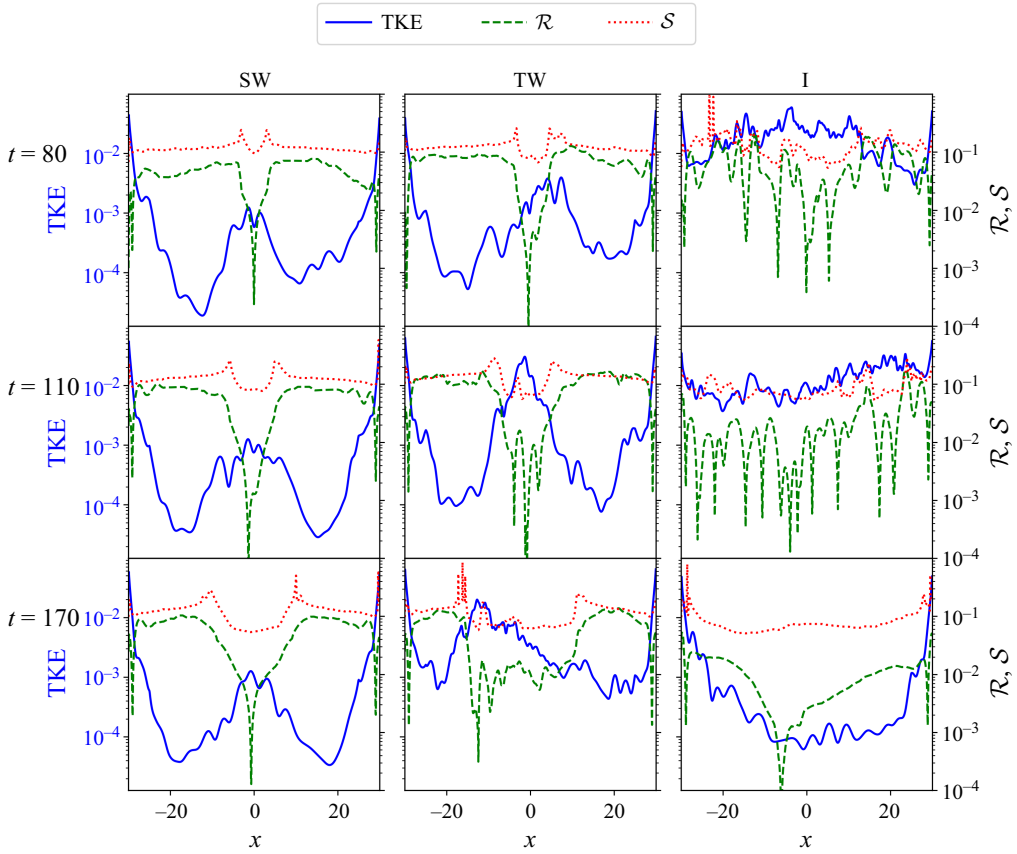


Figure 20. Instantaneous profiles of the TKE from (3.25) averaged over the duct cross-section, along with the long–long wave resonance indicator \mathcal{R} from (4.8) and the long–short wave resonance indicator \mathcal{S} from (4.11), at different times for the SW, TW and I cases.

To quantify resonant interactions in DNS data, we measure the relative spectral gap between perturbed upper and lower layers with respect to the isolated middle layer as

$$\mathcal{R}(\mathbf{W}, \mathbf{Y}) \equiv \frac{|S_{\mathbf{Z}}(\mathbf{W}) - S_{\mathbf{Z}}(\mathbf{Y})|}{S_{\mathbf{Z}}(\mathbf{W}) + S_{\mathbf{Z}}(\mathbf{Y})}, \quad (\text{B14})$$

which yields $\mathcal{R}(\mathbf{W}, \mathbf{Y}) = \mathcal{R}(\mathbf{Y}, \mathbf{W})$, whereas $S_{\mathbf{Z}}(\mathbf{W}) \neq S_{\mathbf{W}}(\mathbf{Z}) \neq S_{\mathbf{Z}}(\mathbf{Y})$ in general. From the definition of $\mathcal{R}(\mathbf{W}, \mathbf{Y})$ in (B14), and employing (B13) to compute $S_{\mathbf{Z}}(\mathbf{W})$ and $S_{\mathbf{Z}}(\mathbf{Y})$ while utilising the fact that $\lambda(\mathbf{Z})$ has only one real value, we deduce (4.8).

Appendix C. Instantaneous profiles of TKE, \mathcal{R} and \mathcal{S}

Figure 20 presents the instantaneous profiles of the TKE, averaged over the duct cross-section, alongside the long–long wave resonance indicator \mathcal{R} and long–short wave resonance indicator \mathcal{S} at $t = 0$, $t = 110$ and $t = 170$ for the SW, TW and I cases. Although the space–time distributions of these quantities are already shown in figures 9, 12 and 15, these profiles provide a more detailed visualisation of their distribution at specific times.

For all cases, \mathcal{R} reaches local minima near the ends of the duct, where the TKE is most pronounced. In contrast, \mathcal{S} is not minimised at the duct ends, suggesting that long–short wave resonance does not occur there; instead, the route to turbulence stems from both

long–long wave resonance and short-wave instability without significant interaction with long waves.

In the wave cases, as we move towards the centre of the duct, long–short wave resonance becomes more prominent, with \mathcal{S} reaching minimum values where small local minima of \mathcal{R} and strong TKE are observed. For the I case, the entire duct exhibits strong fluctuations at $t = 80$ and $t = 110$ due to intense turbulence, indicating that the transition has likely occurred, making further discussion of hydraulic transition less relevant. However, at $t = 170$, fluctuations are significantly reduced due to the intermittent nature of this case. Here, \mathcal{R} shows minimum values near the duct centre, and \mathcal{S} is reduced in a pattern similar to that observed in the wave case. This suggests that both long–long and long–short wave resonances may also occur during the quiet phase of the intermittent case, potentially leading to stronger turbulence at later times.

Appendix D. Information propagation in the presence of complex characteristics

When λ is real, the coordinate transformation $\xi_{\pm} = t \pm x/\lambda$ identifies characteristic curves along which combinations of flow variables are conserved in the homogeneous inviscid limit. We now explore the implications of this transformation when $\lambda \in \mathbb{C}$, and the system is extended into the complex plane using the transformation

$$\tilde{x} = x(1 + \epsilon i), \quad \tilde{t} = t(1 + \epsilon i), \tag{D1}$$

where $\epsilon \ll 1$ is a small parameter and i is the imaginary unit. Using the chain rule we relate spatiotemporal variation of the state vector $\mathbf{q}(x, t)$ and the transformed state vector $\mathbf{q}(\tilde{x}, \tilde{t})$ directly, i.e.

$$\frac{\partial \mathbf{q}(x, t)}{\partial x} = \text{Re} \left((1 + \epsilon i) \frac{\partial \mathbf{q}(\tilde{x}, \tilde{t})}{\partial \tilde{x}} \right), \quad \frac{\partial \mathbf{q}(x, t)}{\partial t} = \text{Re} \left((1 + \epsilon i) \frac{\partial \mathbf{q}(\tilde{x}, \tilde{t})}{\partial \tilde{t}} \right), \tag{D2}$$

to leading order. We also expand $\lambda(x, t)$ in terms of (\tilde{x}, \tilde{t}) using a Taylor series expansion

$$\lambda(x, t) = \lambda(\tilde{x}, \tilde{t}) - \epsilon i \left(\tilde{x} \frac{\partial \lambda(\tilde{x}, \tilde{t})}{\partial \tilde{x}} + \tilde{t} \frac{\partial \lambda(\tilde{x}, \tilde{t})}{\partial \tilde{t}} \right) + \mathcal{O}(\epsilon^2). \tag{D3}$$

Substituting (D2) and (D3) into

$$D_{(x,t)} \mathbf{q}(x, t) = \frac{\partial \mathbf{q}(x, t)}{\partial t} + \lambda(x, t) \frac{\partial \mathbf{q}(x, t)}{\partial x}, \tag{D4}$$

and using (3.8), we obtain the modified equation in terms of (\tilde{x}, \tilde{t}) , i.e.

$$D_{(x,t)} \mathbf{q}(x, t) = \text{Re} \left(\underbrace{\frac{\partial \mathbf{q}(\tilde{x}, \tilde{t})}{\partial \tilde{t}} + \lambda(\tilde{x}, \tilde{t}) \frac{\partial \mathbf{q}(\tilde{x}, \tilde{t})}{\partial \tilde{x}}}_{D_{(\tilde{x}, \tilde{t})} \mathbf{q}(\tilde{x}, \tilde{t})} - \text{Re} \left(\epsilon i \left[\tilde{x} \left(\lambda \frac{\partial^2 \mathbf{q}}{\partial \tilde{x}^2} + \frac{\partial \lambda}{\partial \tilde{x}} \frac{\partial \mathbf{q}}{\partial \tilde{x}} \right) + \tilde{t} \left(\lambda \frac{\partial^2 \mathbf{q}}{\partial \tilde{x} \partial \tilde{t}} + \frac{\partial \lambda}{\partial \tilde{t}} \frac{\partial \mathbf{q}}{\partial \tilde{x}} \right) \right] + \mathcal{O}(\epsilon^2) \right), \tag{D5}$$

assuming that $\mathbf{q}(\tilde{x}, \tilde{t})$ and $\lambda(\tilde{x}, \tilde{t})$ are both analytic.

D.1. Transformation to characteristic variables

We now introduce characteristic coordinates

$$\xi_+ = \tilde{t} + \frac{\tilde{x}}{\lambda(\tilde{x}, \tilde{t})}, \quad \xi_- = \tilde{t} - \frac{\tilde{x}}{\lambda(\tilde{x}, \tilde{t})}. \tag{D6}$$

These define complex-valued curves in space–time. The real parts of these coordinates correspond to directions of propagation in physical space, with inverse

$$\tilde{t} = \frac{1}{2}(\xi_+ + \xi_-), \quad \tilde{x} = \frac{\lambda}{2}(\xi_+ - \xi_-), \tag{D7}$$

and, therefore, become complex if characteristics are complex, which justifies our change of coordinate from x and t to \tilde{x} and \tilde{t} . Using the chain rule

$$\frac{\partial}{\partial \tilde{x}} = \frac{\partial \xi_{\pm}}{\partial \tilde{x}} \frac{d}{d \xi_{\pm}}, \quad \frac{\partial}{\partial \tilde{t}} = \frac{\partial \xi_{\pm}}{\partial \tilde{t}} \frac{d}{d \xi_{\pm}}, \tag{D8}$$

with

$$\begin{aligned} \frac{\partial \xi_+}{\partial \tilde{x}} &= \frac{1}{\lambda} - \frac{\tilde{x}}{\lambda^2} \frac{\partial \lambda}{\partial \tilde{x}}, & \frac{\partial \xi_+}{\partial \tilde{t}} &= 1 - \frac{\tilde{x}}{\lambda^2} \frac{\partial \lambda}{\partial \tilde{t}}, \\ \frac{\partial \xi_-}{\partial \tilde{x}} &= -\frac{1}{\lambda} - \frac{\tilde{x}}{\lambda^2} \frac{\partial \lambda}{\partial \tilde{x}}, & \frac{\partial \xi_-}{\partial \tilde{t}} &= 1 - \frac{\tilde{x}}{\lambda^2} \frac{\partial \lambda}{\partial \tilde{t}}. \end{aligned} \tag{D9}$$

Substituting (D9) into (D5), and applying the chain rule for all first- and second-order derivatives, we obtain

$$D_{(x,t)} \mathbf{q}(x, t) = \text{Re} \left(\underbrace{A_{\pm}(\tilde{x}, \tilde{t}) \frac{d\mathbf{q}(\xi_{\pm})}{d\xi_{\pm}}}_{D_{(\tilde{x}, \tilde{t})} \mathbf{q}(\tilde{x}, \tilde{t})} + \epsilon i \cdot C_{\pm}(\tilde{x}, \tilde{t}, \mathbf{q}(\xi_{\pm})) + \mathcal{O}(\epsilon^2) \right), \tag{D10}$$

where the coefficient functions are

$$A_+(\tilde{x}, \tilde{t}) = \frac{2\lambda^2 - \tilde{x}(\lambda_{\tilde{x}} + \lambda_{\tilde{t}})}{\lambda^2}, \quad A_-(\tilde{x}, \tilde{t}) = -\frac{\tilde{x}(\lambda_{\tilde{x}}\lambda + \lambda_{\tilde{t}})}{\lambda^2}, \tag{D11}$$

where we denote $\lambda_{\tilde{x}} := (\partial \lambda / \partial \tilde{x})$ and $\lambda_{\tilde{t}} := (\partial \lambda / \partial \tilde{t})$. The terms $C_{\pm}(\tilde{x}, \tilde{t}, \mathbf{q})$ denote first-order correction terms that arise from the chain rule and Taylor series expansions, and include all contributions involving second derivatives of \mathbf{q} , as well as spatial and temporal derivatives of λ , evaluated at (\tilde{x}, \tilde{t}) . Therefore, to leading order, we can reconstruct (3.8) in the transformed (\tilde{x}, \tilde{t}) coordinate as

$$\mathbf{v}^H \mathbf{C} (D_{(x,t)} \mathbf{q}(x, t)) = \mathbf{v}^H \mathbf{C} \text{Re} \left(D_{(\tilde{x}, \tilde{t})} \mathbf{q}(\tilde{x}, \tilde{t}) \right) = \mathbf{0}. \tag{D12}$$

In the limit of weak spatiotemporal variation of λ , i.e. $\lambda_{\tilde{x}}, \lambda_{\tilde{t}} \approx 0$, we find that A_{\pm} become constant, yielding

$$\mathbf{v}^H \mathbf{C} \frac{d\mathbf{q}(\xi_{\pm}^R)}{d\xi_{\pm}^R} = \mathbf{0}. \tag{D13}$$

Thus, a combination of flow variables in the transformed coordinate is conserved to leading order along the real component of the complex characteristics. Hence, to leading order, information propagates along the complex characteristic curve ξ_{\pm} , with the direction

governed approximately by the real part λ^R . Therefore, even in the complex setting, the effective characteristic curve in real space–time is given by

$$\xi_+^R = t + \frac{x}{\lambda^R}, \quad \xi_-^R = t - \frac{x}{\lambda^R}, \quad (\text{D14})$$

which reflects the fact that information propagation along complex characteristics on the leading order is tied to λ^R .

REFERENCES

- ARMI, L. 1986 The hydraulics of two flowing layers with different densities. *J. Fluid Mech.* **163**, 27–58.
- ATOUI, A., ZHU, L., LEFAUVE, A., TAYLOR, J.R., KERSWELL, R.R., DALZIEL, S.B., LAWRENCE, G.A. & LINDEN, P. 2023 Stratified inclined duct: two-layer hydraulics and instabilities. *J. Fluid Mech.* **977**, A25.
- BAINES, P.G. 1988 A general method for determining upstream effects in stratified flow of finite depth over long two-dimensional obstacles. *J. Fluid Mech.* **188**, 1–22.
- BARTHOLOMEW, P., DESKOS, G., FRANTZ, R.A.S., SCHUCH, F.N., LAMBALLAIS, E. & LAIZET, S. 2020 Xcompact3D: An open-source framework for solving turbulence problems on a Cartesian mesh. *SoftwareX* **12**, 100550.
- BENTON, G.S. 1954 The occurrence of critical flow and hydraulic jumps in a multi-layered fluid system. *J. Atmos. Sci.* **11** (2), 139–150.
- BRAY, N.A., OCHOA, J. & KINDER, T.H. 1995 The role of the interface in exchange through the Strait of Gibraltar. *J. Geophys. Res.: Oceans* **100** (C6), 10755–10776.
- BURCHARD, H., LANGE, X., KLINGBEIL, K. & MACCREADY, P. 2019 Mixing estimates for estuaries. *J. Phys. Oceanogr.* **49** (2), 631–648.
- CAULFIELD, C. 2020 Open questions in turbulent stratified mixing: do we even know what we do not know? *Phys. Rev. Fluids* **5** (11), 110518.
- CHANT, R.J. & WILSON, R.E. 2000 Internal hydraulics and mixing in a highly stratified estuary. *J. Geophys. Res.: Oceans* **105** (C6), 14215–14222.
- DALZIEL, S.B. 1991 Two-layer hydraulics: a functional approach. *J. Fluid Mech.* **223**, 135–163.
- DUCIMETIÈRE, Y.M., GALLAIRE, F., LEFAUVE, A. & CAULFIELD, C.P. 2021 Effects of spanwise confinement on stratified shear instabilities. *Phys. Rev. Fluids* **6** (10), 103901.
- ECKART, C. 1961 Internal waves in the ocean. *Phys. Fluids* **4** (7), 791–799.
- ELSNER, L. & SUN, J. 1982 Perturbation theorems for the generalized eigenvalue problem. *Linear Algebr. Applics.* **48**, 341–357.
- ENGQVIST, A. 1996 Self-similar multi-layer exchange flow through a contraction. *J. Fluid Mech.* **328**, 49–66.
- FARMER, D.M. & ARMI, L. 1988 The flow of Atlantic water through the Strait of Gibraltar. *Prog. Oceanogr.* **21** (1), 1–103.
- GALLMEIER, F., SU, Z., ROCHA, B.C., MENEMENLIS, D. & KLEIN, P. 2023 An evaluation of the LLC4320 global ocean simulation based on the submesoscale structure of modeled sea surface temperature fields. *Geosci. Model Dev.* **16** (23), 7143–7164.
- GEYER, W.R., SCULLY, M.E. & RALSTON, D.K. 2008 Quantifying vertical mixing in estuaries. *Environ. Fluid Mech.* **8** (5–6), 495–509.
- GILL, A.E. 1977 The hydraulics of rotating-channel flow. *J. Fluid Mech.* **80** (4), 641–671.
- JAROSZ, E., TEAGUE, W.J., BOOK, J.W. & BEŞİKTEPE, Ş.T. 2012 Observations on the characteristics of the exchange flow in the Dardanelles Strait. *J. Geophys. Res.* **117** (C11012), 1–18.
- JIANG, X., ATOUI, A., ZHU, L., LEFAUVE, A., TAYLOR, J.R., DALZIEL, S.B. & LINDEN, P.F. 2023 Geometry of stratified turbulent mixing: local alignment of the density gradient with rotation, shear and viscous dissipation. *J. Fluid Mech.* **977**, R5.
- JIANG, X., LEFAUVE, A., DALZIEL, S.B. & LINDEN, P. 2022 The evolution of coherent vortical structures in increasingly turbulent stratified shear layers. *J. Fluid Mech.* **947**, A30.
- LANE-SERFF, G.F., SMEED, D.A. & POSTLETHWAITE, C.R. 2000 Multi-layer hydraulic exchange flows. *J. Fluid Mech.* **416**, 269–296.
- LAWRENCE, G.A. 1990 On the hydraulics of Boussinesq and non-Boussinesq two-layer flows. *J. Fluid Mech.* **215**, 457–480.
- LAWRENCE, G.A. 1993 The hydraulics of steady two-layer flow over a fixed obstacle. *J. Fluid Mech.* **254**, 605–633.
- LEFAUVE, A. & LINDEN, P. 2020 Buoyancy-driven exchange flows in inclined ducts. *J. Fluid Mech.* **893**, A2.

A. Atoufi and others

- LEFAUVE, A., PARTRIDGE, J., ZHOU, Q., DALZIEL, S., CAULFIELD, C. & LINDEN, P. 2018 The structure and origin of confined Holmboe waves. *J. Fluid Mech.* **848**, 508–544.
- LONG, R. 1956 Long waves in a two-fluid system. *J. Atmos. Sci.* **13** (1), 70–74.
- MA, Y. 1981 The resonant interaction among long and short waves. *Wave Motion* **3** (3), 257–267.
- MEYER, C. & LINDEN, P. 2014 Stratified shear flow: experiments in an inclined duct. *J. Fluid Mech.* **753**, 242–253.
- PRATT, L.J., JOHNS, W., MURRAY, S.P. & KATSUMATA, K. 1999 Hydraulic interpretation of direct velocity measurements in the Bab al Mandab. *J. Phys. Oceanogr.* **29** (11), 2769–2784.
- SANNINO, G., CARILLO, A. & ARTALE, V. 2007 Three-layer view of transports and hydraulics in the Strait of Gibraltar: a three-dimensional model study. *J. Geophys. Res.* **112** (C03010), 1–20.
- SANNINO, G., PRATT, L. & CARILLO, A. 2009 Hydraulic criticality of the exchange flow through the Strait of Gibraltar. *J. Phys. Oceanogr.* **39** (11), 2779–2799.
- SHINNERS, S.M. 1998 *Modern Control System Theory and Design*. John Wiley & Sons.
- SMEED, D.A. 2000 Hydraulic control of three-layer exchange flows: application to the Bab el Mandab. *J. Phys. Oceanogr.* **30** (10), 2574–2588.
- SMYTH, W.D. & CARPENTER, J.R. 2019 *Instability in Geophysical Flows*. Cambridge University Press.
- SMYTH, W.D. & PELTIER, W.R. 1991 Instability and transition in finite-amplitude Kelvin–Helmholtz and Holmboe waves. *J. Fluid Mech.* **228**, 387–415.
- STEWART, G.W. 1975 Gershgorin theory for the generalized eigenvalue problem $Ax = \lambda Bx$. *Math. Comput.* **29** (130), 600–606.
- STEWART, G.W. & SUN, J.G. 1990 *Matrix Perturbation Theory*. Academic Press.
- SÁNCHEZ-GARRIDO, J.C., SANNINO, G., LIBERTI, L., GARCÍA LAFUENTE, J. & PRATT, L. 2011 Numerical modeling of three-dimensional stratified tidal flow over Camarinal Sill, Strait of Gibraltar. *J. Geophys. Res.* **116** (C12026), 1–17.
- ZHU, L., ATOUFI, A., LEFAUVE, A., KERSWELL, R.R. & LINDEN, P.F. 2024 Long-wave instabilities of sloping stratified exchange flows. *J. Fluid Mech.* **983**, A12.
- ZHU, L., ATOUFI, A., LEFAUVE, A., TAYLOR, J.R., KERSWELL, R.R., DALZIEL, S.B., LAWRENCE, G.A. & LINDEN, P. 2023 Stratified inclined duct: direct numerical simulations. *J. Fluid Mech.* **969**, A20.