On the Rigid-Lid Approximation for Two Shallow Layers of Immiscible Fluids with Small Density Contrast

Vincent Duchêne

Journal of Nonlinear Science

ISSN 0938-8974
Volume 24
Number 4

J Nonlinear Sci (2014) 24:579-632
DOI 10.1007/s00332-014-9200-2

Springer

J Nonlinear Sci (2014) 24:579–632
DOI 10.1007/s00332-014-9200-2

Journal of Nonlinear Science

On the Rigid-Lid Approximation for Two Shallow Layers of Immiscible Fluids with Small Density Contrast

Vincent Duchêne

Received: 11 October 2013 / Accepted: 6 February 2014 / Published online: 28 March 2014

Abstract The rigid-lid approximation is a commonly used simplification in the study of density-stratified fluids in oceanography. Roughly speaking, one assumes that the displacements of the surface are negligible compared with interface displacements. In this paper, we offer a rigorous justification of this approximation in the case of two shallow layers of immiscible fluids with constant and quasi-equal mass density. More precisely, we control the difference between the solutions of the Cauchy problem predicted by the shallow-water (Saint-Venant) system in the rigid-lid and free-surface configuration. We show that in the limit of a small density contrast, the flow may be accurately described as the superposition of a baroclinic (or slow) mode, which is well predicted by the rigid-lid approximation, and a barotropic (or fast) mode, whose initial smallness persists for large time. We also describe explicitly the first-order behavior of the deformation of the surface and discuss the case of a nonsmall initial barotropic mode.

Keywords Internal waves · Rigid-lid approximation · Boussinesq approximation · Asymptotic analysis · Hyperbolic system

1 Introduction

1.1 Motivation

The mass density of water in the ocean is not constant due to variations in temperature and salinity. As a matter of fact, one typically observes a sharp separation between a layer of warm, relatively fresh water above a layer of cold, more salty water. The

Communicated by P. Newton.

V. Duchêne (✉)
Université de Rennes, Institut de Recherche Mathématique de Rennes, France
e-mail: vincent.duchene@univ-rennes1.fr

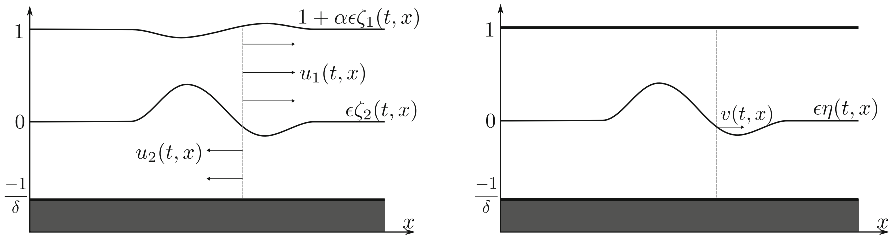
interface between these two layers may experience great deformations that are mostly invisible at the surface but account for important oceanographic features, such as internal solitary waves or the dead-water phenomenon (e.g., Gill 1982; Jackson 2004; Helfrich and Melville 2006 and references therein). The study of these internal waves has attracted a considerable amount of attention in recent decades and led to a vast collection of various models. To simplify the setting, two approximations are commonly used in the literature—the rigid-lid and Boussinesq approximations. Roughly speaking, the rigid-lid approximation consists in neglecting the surface displacements compared to interface displacements, while the Boussinesq approximation relies on the assumption that the density differences between the two layers is small. Admittedly, these two assumptions are related: a fixed amount of energy generates a much smaller displacement at the air/water interface than at the fresh/salty water interface because the ratio of mass densities across the interface is negligible in the former case compared to the latter.

The ambition of this article is to offer a rigorous justification of the foregoing presumption. We restrict ourselves to one of the simplest possible settings, that is, two infinite, two-dimensional layers of immiscible fluids with constant density above a flat bottom. Moreover, we consider sufficiently shallow layers so that the hydrostatic approximation is valid; thus we study the so-called Saint-Venant (1871) or shallow-water equations. Even in that much simplified setting, we will come across serious difficulties that derive from the fact that the typical surface wave speed, as predicted by the linearized system, is much greater than the typical interface wave speed, in particular in the limit of a vanishing density contrast. Thus, within the terms neglected in the rigid-lid approximation are contributions whose velocity blows up in the limit we consider. As a matter of fact, even the well-posedness of the Cauchy problem for the Saint-Venant system in the free-surface configuration on a relevant time scale (i.e., nonvanishing with the density contrast) is challenging.

To our knowledge, very few works have been concerned with the validity of the aforementioned approximations, despite the early concerns expressed by Long (1965) and Benjamin (1966). Grimshaw et al. (2002), Craig et al. (2005), Craig et al. (2010), and Duchêne (2010) derive and compare asymptotic models in both the rigid-lid and free-surface settings. However, they do not directly compare solutions of the two models with corresponding initial data but rather parameters of their models, or explicit solutions (solitary waves). Moreover, and perhaps more importantly, their analysis is restricted to weakly nonlinear waves, so that the deformation of both the surface and interface is assumed to be small. Recently, Leonardi (2011) studied in much detail the validity of the rigid-lid approximation in a linearized setting and without explicitly looking at the limit of small density differences. Conversely, our study accounts for fully nonlinear waves and directly compares the solutions predicted by the rigid-lid and free-surface systems in the limit of a vanishing density contrast.

1.2 Presentation of Models, and Main Result

In this section, we present the two models we study—the shallow-water (or Saint-Venant) system in free-surface and rigid-lid configurations (Fig. 1). We briefly describe



(a) Free-surface situation **(b)** Rigid-lid situation

Fig. 1 Sketch of the domain in the two different situations at stake

some early properties of these models and state our main result in Theorem 1.2. There follows an outline of the present paper and some notations used therein.

Free-surface system. Let us first introduce the shallow-water model with a free surface, which we simply refer to as the *free-surface system*:

$$\begin{cases} \alpha\partial_t\zeta_1 + \partial_x(h_1u_1) + \partial_x(h_2u_2) = 0, \\ \partial_t\zeta_2 + \partial_x(h_2u_2) = 0, \\ \partial_t u_1 + \alpha\frac{\delta+\gamma}{1-\gamma}\partial_x\zeta_1 + \frac{\epsilon}{2}\partial_x\left(|u_1|^2\right) = 0, \\ \partial_t u_2 + (\delta+\gamma)\partial_x\zeta_2 + \gamma\alpha\frac{\delta+\gamma}{1-\gamma}\partial_x\zeta_1 + \frac{\epsilon}{2}\partial_x\left(|u_2|^2\right) = 0, \end{cases} \tag{1.1}$$

where we use the notation $h_1 = 1 + \epsilon\alpha\zeta_1 - \epsilon\zeta_2$ and $h_2 = \frac{1}{\delta} + \epsilon\zeta_2$.

This system was obtained[1] in Choi and Camassa (1996), Craig et al. (2005) and justified in Duchêne (2010) as an asymptotic model (in the shallow-water regime) for a system of two layers of immiscible, homogeneous, ideal, incompressible fluid under the sole influence of gravity (the so-called full Euler system). It describes the evolution of the deformation of the surface, ζ_1, the interface, ζ_2, and the horizontal velocity of the fluid in the upper (resp. lower) layer, u_1 (resp. u_2).[2] More precisely, the two layers are assumed to be connected, infinite in the horizontal dimension $x \in \mathbb{R}$, delimited below by a flat bottom and by the graph of functions $\zeta_1(t,x)$, $\zeta_2(t,x)$ (Fig. 1).

The parameters $\alpha, \delta, \gamma, \epsilon$ are dimensionless parameters that describe the characteristics of the flow. More precisely:

- δ represents the ratio of the upper-layer to the lower-layer depth;
- γ represents the ratio of the mass density between the two fluids;
- ϵ represents the maximal deformation of the interface divided by the upper-layer depth;

[1] The models presented in these works are not limited to a flat bottom or horizontal dimension $d = 1$. They present different constants in the velocity equations. This is due to a different choice of scaling in the nondimensionalizing step. We chose our scaling in order to set the typical velocity of the internal wave (obtained by solving explicitly the linear system, i.e., setting $\alpha = \epsilon = 0$) as $c_0 = \pm 1$, consistently with the rigid-lid system (1.2).

[2] The Saint-Venant model is usually derived using the so-called hydrostatic approximation. Equivalently, one may assume that the horizontal scale is large compared with the vertical scale, so that the horizontal velocity field is accurately described as constant throughout the depth of each layer of fluid.

α represents the ratio of the maximal deformation of the surface to that of the interface.

In particular, h_1 denotes the depth of the upper layer and h_2 the depth of the lower layer.

Remark 1.1 Another dimensionless parameter plays an important role but is not visible here, although it is essential for the construction and relevance of the shallow-water models. If we denote by μ the ratio of the depth of the two layers to a characteristic horizontal length, then one assumes $\mu \ll 1$, and all terms of size $\mathcal{O}(\mu^2)$ are neglected in (1.1).

An additional dimensionless parameter is ubiquitous in the present work and obtained as a combination of the aforementioned parameters. It turns out to be convenient to express the assumption that the density contrast between the two fluids is small with

$$\varrho \ll 1 \quad ; \quad \varrho \equiv \sqrt{\frac{1-\gamma}{\gamma+\delta}}.$$

We conclude the presentation of the free-surface system by mentioning that system (1.1) is obviously a system of four conservation laws but also induces at least two other conserved quantities. Indeed, as noticed in Barros et al. (2007), after manipulating the equations, one may obtain the following identities:

- Conservation of horizontal momentum:

$$\partial_t(\gamma h_1 u_1 + h_2 u_2) + \partial_x p + \partial_x(\gamma h_1 |u_1|^2 + h_2 |u_2|^2) = 0,$$

where p is the "pressure": $p = \frac{1}{2}\left(\gamma\frac{\delta+\gamma}{1-\gamma}(h_1+h_2)^2 + (\gamma+\delta)h_2^2\right)$.
- Conservation of energy:

$$\partial_t E + \partial_x \left(\frac{1}{2}(\gamma h_1 |u_1|^2 u_1 + h_2 |u_2|^2 u_2) + \gamma h_1^2 u_1 + h_2^2 u_2 + \gamma h_1 h_2 (u_1+u_2)\right) = 0,$$

where we use the notation $E \equiv \frac{1}{2}\gamma h_1 |u_1|^2 + \frac{1}{2} h_2 |u_2|^2 + p$.

Rigid-lid system. The model corresponding to (1.1) in the rigid-lid configuration, which we refer to as the *rigid-lid system*, is

$$\begin{cases} \partial_t \eta \;+\; \partial_x\left(\frac{h_1 h_2}{h_1+\gamma h_2} v\right) \;=\; 0, \\ \partial_t v \;+\; (\gamma+\delta)\partial_x \eta \;+\; \frac{\epsilon}{2}\partial_x\left(\frac{h_1^2-\gamma h_2^2}{(h_1+\gamma h_2)^2}|v|^2\right) \;=\; 0. \end{cases} \tag{1.2}$$

Here, η represents the deformation of the interface and v the shear velocity, namely, $v = u_2 - \gamma u_1$; see below and Fig. 1. Again, h_1 and h_2 denote the depth of the upper (resp. lower) layers, and thus $h_1 = 1 - \epsilon\eta$ and $h_2 = 1/\delta + \epsilon\eta$. The parameters γ, δ, ϵ are defined as previously.

System (1.2) has been justified as an asymptotic model in the shallow-water regime in Bona et al. (2008),[3] starting from the full Euler system in the rigid-lid configuration. Let us show how to *formally* recover (1.2) from (1.1). Set $\zeta_1 \equiv 0$ (or, equivalently, $\alpha = 0$) in (1.1). It follows in particular from the first equation that

$$\partial_x(h_1 u_1) + \partial_x(h_2 u_2) \;=\; 0. \tag{1.3}$$

Since $h_1 u_1$ and $h_2 u_2$ are scalar functions vanishing at infinity, we deduce the identity $h_1 u_1 = -h_2 u_2$. Thus, when we define $v \equiv u_2 - \gamma u_1$, we obtain

$$u_1 \equiv \frac{-h_2 v}{h_1 + \gamma h_2} \quad \text{and} \quad u_2 \equiv \frac{h_1 v}{h_1 + \gamma h_2}. \tag{1.4}$$

It is now clear that the second equation and a linear combination of the last two equations of (1.1) yield (1.2) (with $\eta \equiv \zeta_2$). We aim to give a rigorous confirmation of the preceding calculations.

Main result. We state here the main result of the present work.

Theorem 1.2 *Let $s \geq s_0 + 1, s_0 > 1/2$, and $\delta_{\min}, \delta_{\max}, \gamma_{\min} > 0$. Consider $(\alpha, \delta, \epsilon, \gamma) \in \mathcal{P}$, with*

$$\mathcal{P} \equiv \Big\{(\alpha, \delta, \epsilon, \gamma),\; 0 \leq \alpha \leq 1, \quad \delta_{\min} \leq \delta \leq \delta_{\max}, \quad 0 < \epsilon \leq 1, \\ \gamma_{\min} \leq \gamma < 1\Big\}.$$

Let $\zeta_1^0, \zeta_2^0, u_1^0, u_2^0 \in H^{s+1}(\mathbb{R})$ satisfy the following hypotheses:

$$\begin{aligned} &\big|\zeta_2^0\big|_{H^{s+1}} + \big|u_2^0 - \gamma u_1^0\big|_{H^{s+1}} \leq M \quad \textit{and} \\ &\frac{\alpha}{\varrho}\big|\zeta_1^0\big|_{H^{s+1}} + \big|\gamma h_1 u_1^0 + h_2 u_2^0\big|_{H^{s+1}} \leq M\,\varrho, \end{aligned} \tag{1.5}$$

as well as (denoting $h_1^0 \equiv 1 + \epsilon\alpha\zeta_1^0 - \epsilon\zeta_2^0$ and $h_2^0 \equiv \delta^{-1} + \epsilon\zeta_2^0$)

$$\forall x \in \mathbb{R}, \quad \min\Big\{h_1^0(x)\;;\; h_2^0(x) - \epsilon^2\frac{|u_2^0(x) - u_1^0(x)|^2}{\gamma + \delta}\;; \\ (h_1^0(x) + \gamma h_2^0(x))^3 - \epsilon^2\frac{\gamma(1 + \delta^{-1})^2|u_2^0(x) - \gamma u_1^0(x)|^2}{\gamma + \delta}\Big\} \geq h_0 > 0, \tag{1.6}$$

[3] The justification provided in Bona et al. (2008)—as well as in Duchêne (2010) in the free-surface configuration—is in the sense of consistency: sufficiently smooth solutions of the full Euler system satisfy the equations of (1.2) up to small, i.e., $\mathcal{O}(\mu^2)$, remainder terms. The rigorous, full justification follows from the well-posedness of both the full Euler system and the shallow-water model, as well as a stability result, which make it possible to compare the solutions of both systems with corresponding initial data on the relevant time scale. In the rigid-lid situation, Lannes (2013) recently solved the difficult problem of the well-posedness of the full Euler system, consequently completing the full justification of (1.2); see Lannes (2013), Theorem 7. No such result is available in the bifluidic, free-surface configuration.

where $0 < h_0, M < \infty$ *are fixed.*

Then there exist T^{-1}, C*, positive, depending only and nondecreasingly on* $M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}$ *and* $\frac{1}{s_0-\frac{1}{2}}$*, such that the following assertions hold.*

1. *There exists a unique solution,* $(\eta, v) \in C([0, T/(\epsilon M)]; H^{s+1}(\mathbb{R})^2) \cap C^1([0, T/(\epsilon M)]; H^s(\mathbb{R})^2)$ *to* (1.2)*, with initial data* $(\eta\,|_{t=0} = \zeta_2^0, v\,|_{t=0} = u_2^0 - \gamma u_1^0)$.
2. *There exists a unique solution,* $(\zeta_1, \zeta_2, u_1, u_2) \in C([0, T_{\max}); H^{s+1}(\mathbb{R})^4) \cap C^1([0, T_{\max}); H^s(\mathbb{R})^4)$ *to* (1.1)*, with initial data* $(\zeta_1^0, \zeta_2^0, u_1^0, u_2^0)$*, and* $T_{\max} \geq T/\max\{\epsilon M, \varrho\}$.
3. *One has, for any* $0 \leq t \leq T/\max\{\epsilon M, \varrho\}$,

$$\frac{\alpha}{\varrho}\big\|\zeta_1\big\|_{L^\infty([0,t];H^s)} + \big\|\gamma h_1 u_1 + h_2 u_2\big\|_{L^\infty([0,t];H^s)} \leq C\,M\,\varrho$$

and

$$\big\|\eta - \zeta_2\big\|_{L^\infty([0,t];H^s)} + \big\|v - (u_2 - \gamma u_1)\big\|_{L^\infty([0,t];H^s)} \leq C\,M\,\varrho.$$

Remark 1.3 The restriction on the maximal time of existence for the solution of the free-surface system, $T_{\max} \geq T/\max\{\epsilon M, \varrho\}$, as opposed to the classical $T_{\max} \geq T/(\epsilon M)$, is purely technical and does not reveal any limitation that would appear in the weakly nonlinear case, $\epsilon M = \mathcal{O}(\varrho)$. In contrast, we know that in the latter case (Proposition 2.2 and Remark 2.4), system (1.1) is well-posed over time $T_{\max} \gtrsim (\epsilon M)^{-1}$, without the additional condition in (1.5). Moreover, it would not be difficult to obtain an asymptotic description of the solution similar to that obtained by Duchêne (2010) (without the dispersion terms), namely that the flow may be accurately approximated as a superposition of four independent waves, each driven by an inviscid Burgers equation. The solution of the rigid-lid system (1.2) conforms to a similar description (with only two counterpropagating waves); thus the two solutions are easily compared. We present in Sect. 4 a similar decomposition of the flow allowing for stronger nonlinearities; see in particular Theorem 4.5 and Proposition 4.6.

To acknowledge the fact that we are interested in strong nonlinearities, and to facilitate reading, *we set* $\epsilon \equiv 1$ in what follows.

Remark 1.4 The factor $\frac{\alpha}{\varrho}$ in front of ζ_1 is natural in our context. Indeed, one easily deduces from the aforementioned conservation of energy for (1.1) that

$$\int_{\mathbb{R}} E(x) - E(\infty)\,dx \approx \frac{\gamma}{\varrho^2}\big|\alpha\zeta_1\big|_{L^2}^2 + \big|\zeta_2\big|_{L^2}^2 + \gamma\big|u_1\big|_{L^2}^2 + \big|u_2\big|_{L^2}^2$$

is constant in time, so that without any further assumption than a finite initial energy we know that $\gamma^{1/2}\frac{\alpha}{\varrho}\big|\zeta_1\big|_{L^2}$ remains bounded as long as the solution is well defined. For simplicity's sake, *we set* $\alpha \equiv \varrho$ in what follows.

Let us emphasize again the consequences of the assumptions made on the preceding remarks. The set of parameters we consider throughout the rest of the paper is

$$\mathcal{P} \equiv \left\{ (\alpha, \delta, \epsilon, \gamma), \quad \alpha = \varrho \equiv \sqrt{\frac{1-\gamma}{\gamma+\delta}}, \quad \delta_{\min} \le \delta \le \delta_{\max}, \quad \epsilon = 1, 0 < \gamma < 1 \right\},$$

with fixed $0 < \delta_{\min} \le \delta_{\max} < \infty$. The interesting limit is therefore $\varrho \to 0$ or, equivalently, $\gamma \to 1$. Except for Sect. 2 and the appendix, we additionally impose $0 < \gamma_{\min} \le \gamma$, with $\gamma_{\min}$ fixed. The assumptions $\epsilon = 1$ and $\alpha = \varrho$ do not lack in generality because one can recover the general case, and in particular the set of parameters in the statement of Theorem 1.2, after applying straightforward scaling factors on the unknowns.

Remark 1.5 Notice that we do not impose any smallness on the parameter ϱ. Of course, for nonsmall ϱ, our result does not improve already existing results in the literature, namely the well-posedness of the Cauchy problem in Sobolev spaces for free-surface and rigid-lid systems (Sect. 2). In that case, one does not expect the free-surface solution to be accurately described by the rigid-lid solution. In other words, *the rigid-lid approximation is not valid if ϱ is not small*; see, for example, the discussion and numerical simulations in Duchêne (2011). When ϱ is small, the essential assumption is the second inequality in (1.5), which can be viewed as an assumption of *well-prepared initial data*: it ensures that the time derivative of the flow is initially bounded, uniformly for ϱ small. Such assumptions are standard in the analysis of singularly perturbed systems; see, e.g., Klainerman and Majda (1981), Browning and Kreiss (1982).

Remark 1.6 A natural extension of our work would consist in treating the situation of the horizontal dimension $d = 2$. The free-surface system in that case would have the same quasilinear structure as (1.1), and a symmetrizer was exhibited in Duchêne (2010). In contrast, the rigid-lid system as constructed in Bona et al. (2008) is quite different as it involves a nonlocal operator constructed from the orthogonal projector onto the gradient vector fields of $L^2(\mathbb{R}^d)^d$. This can be seen from the fact that Eq. (1.3), imposed by the rigid-lid hypothesis, becomes $\nabla \cdot (h_1\mathbf{u}_1 + h_2\mathbf{u}_2) = 0$, which does not enforce $h_1\mathbf{u}_1 + h_2\mathbf{u}_2 = \mathbf{0}$ when $\mathbf{u}_1, \mathbf{u}_2$ map $\mathbb{R}^2$ to $\mathbb{R}^2$ (and, in particular, (1.4) does not hold in general). Let us note, however, that the well-posedness of the shallow-water system in the rigid-lid configuration when $d = 2$ was established in Guyenne et al. (2010), Bresch and Renardy (2011). Interestingly, the system considered in Bresch and Renardy (2011), which is formulated differently than in Bona et al. (2008), Guyenne et al. (2010) and admits nonirrotational velocity fields, offers a clear approximate solution (in the sense of consistency) to the Saint-Venant system in the free-surface configuration.

Remark 1.7 The case of a (sufficiently regular) nonflat bottom topography can be treated following the strategy of this work, after straightforward arrangements. Indeed, the hyperbolic structure of systems (1.1) and (1.2) is not altered when the topography is taken into account, and the only modification is the appearance of a "source" term of the form $\mathbf{f}(U)\partial_x b$, where $\mathbf{f}$ is a vector-valued function depending only on U, the unknown vector-field, and b is the bottom topography. Note, however, that the decomposition between the fast and slow modes introduced in Sect. 4 would not be valid because the persistence of spatial localization (e.g., Lemma 4.4) would not hold with the additional source term.

Remark 1.8 Contrary to the shallow-water systems (1.1) and (1.2), the corresponding full Euler system is ill-posed in Sobolev spaces in the absence of surface (or interface) tension due to the so-called Kelvin–Helmholtz instabilities. In Lannes (2013), Lannes shows that, at least in the rigid-lid configuration, a small amount of interface tension may be sufficient to regularize the high-frequency component of the flow, hence ensuring the existence and uniqueness of a solution to the initial-value problem over large times. By selecting the low-frequency component of the flow, the shallow-water assumption tames the Kelvin–Helmholtz instabilities and allows for our systems to be well-posed even without the corresponding surface tension components. Conditions (1.6), or more precisely the restrictions on the magnitude of the shear velocity that define the domain of hyperbolicity of systems (1.1) and (1.2), are reminiscent of these instabilities.

Outline of the paper. Section 2 is dedicated to some preliminary results on the Cauchy problem for systems (1.1) and (1.2), obtained through classical techniques on quasilinear, hyperbolic systems. Indeed, one easily checks that systems (1.1) and (1.2) are Friedrichs-symmetrizable under reasonable assumptions on the data. As a matter of fact, the Cauchy problem for (1.2) was studied in detail in Guyenne et al. (2010); Bresch and Renardy (2011) (with the much more difficult case of horizontal dimension $d = 2$), and we recall their result in Proposition 2.1.

In the same way, one easily obtains the well-posedness of the Cauchy problem for the free-surface system (1.1) through standard energy methods; we state the result in Proposition 2.2, and postpone its proof to the appendix. However, the resulting time of existence is only of size $T \gtrsim \varrho$. One objective of our work is to obtain a control of the energy over a large time scale (i.e., uniform with respect to ϱ small) and describe the asymptotic behavior of the solution when ϱ vanishes.

Note that Proposition 2.2 also contains the usual blow-up criterion, so that item 2 in Theorem 1.2 is a consequence of the control of the solution on the relevant time scale. Thus it suffices to prove item 3, and the entire statement follows. Section 3 is dedicated to the proof of item 3.

Finally, in Sect. 4, we discuss several natural developments around Theorem 1.2:

- The construction of a first-order corrector term in order to obtain greater precision. In particular, we describe the asymptotic behavior of a small deformation at the surface.
- The case of ill-prepared initial data, that is, data failing to meet the smallness assumption in (1.5).

On both counts, the relevant notion lies in a decomposition between the fast and slow modes (or barotropic and baroclinic modes), which we define precisely subsequently.

Finally, Sect. 4.3 also contains a discussion on the various results of the present work, supported with numerical simulations.

Notations. If not specified, C_0 denotes a nonnegative constant whose exact expression is of no importance. In the present work, C_0 almost always depends nondecreasingly on $\delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}$ and often on $\frac{1}{s_0-1/2}$, such dependency being nonnecessarily specified. The notation $a \lesssim b$ or $a = \mathcal{O}(b)$ means $a \leq C_0 b$, and $a \approx b$ means $a \lesssim b$ and $b \lesssim a$, while $a \sim b$ means $\frac{a}{b} \to 1$ ($\varrho \to 0$).

We denote by $C(\lambda_1, \lambda_2, \ldots)$ a nonnegative constant depending on the parameters $\lambda_1, \lambda_2, \ldots$, and whose dependence on the λ_j is always assumed to be nondecreasing.

The real inner product of any functions f_1 and f_2 in the Hilbert space of square-integrable functions, $L^2 = L^2(\mathbb{R})$, is denoted by

$$\big(f_1 \,,\; f_2 \big) = \int_{\mathbb{R}} f_1(x) f_2(x)\, dx.$$

The space $L^\infty = L^\infty(\mathbb{R})$ consists of all essentially bounded, Lebesgue-measurable functions f, and

$$\big| f \big|_{L^\infty} = \operatorname*{ess\,sup}_{x\in\mathbb{R}} |f(x)| \;<\; \infty.$$

For any real $s \geq 0$, $H^s = H^s(\mathbb{R})$ denotes the Sobolev space of all tempered distributions, f, endowed with the norm $|f|_{H^s} = |\Lambda^s f|_{L^2} < \infty$, where Λ is the fractional derivative $\Lambda = (\mathrm{Id} - \partial_x^2)^{1/2}$.

For any $U \equiv (\zeta_1, \zeta_2, u_1, u_2)^\top \in H^s(\mathbb{R})^4$ and $0 < \gamma < 1$, we introduce the following norm:

$$\big|U\big|_{X^s}^2 \;=\; \gamma\big|\zeta_1\big|_{H^s}^2 + \big|\zeta_2\big|_{H^s}^2 + \gamma\big|u_1\big|_{H^s}^2 + \big|u_2\big|_{H^s}^2.$$

Except in Sect. 2 and the appendix, we assume that γ is uniformly bounded from below, so that X^s is equivalent to the standard $H^s(\mathbb{R})^4$-norm.

For any functions $u = u(t,x)$ and $v(t,x)$ defined on $[0,T) \times \mathbb{R}$ with some $T > 0$, we denote the inner product, the L^2-norm, and the Sobolev norms with respect to the spatial variable x, with $\big(u, v\big) = \big(u(t,\cdot), v(t,\cdot)\big)$, $\big|u\big|_{L^2} = \big|u(t,\cdot)\big|_{L^2}$, and $|u|_{H^s} = |u(t,\cdot)|_{H^s}$, respectively.

For $T > 0$ and X a functional space, we denote by $L^\infty([0,T); X)$ the space of functions such that $u(t,\cdot)$ is controlled in X, uniformly for $t \in [0,T)$. This space is endowed with the following norm:

$$\big\|u\big\|_{L^\infty([0,T);X)} \;=\; \operatorname*{ess\,sup}_{t\in[0,T)} |u(t,\cdot)|_X \;<\; \infty.$$

Finally, $C^k([0,T); X)$ denotes the space of k-times continuously differentiable functions in X.

2 Preliminary Results

In this section, we present some results concerning the Cauchy problem related to the free-surface and rigid-lid systems, respectively (1.1) and (1.2), in Sobolev spaces.

Proposition 2.1 *(Well-posedness result concerning the rigid-lid system)*
Let $s \geq s_0 + 1$, $s_0 > 1/2$, and let $U^0 = (\zeta^0, v^0)^\top \in H^s(\mathbb{R})^2$ be such that there exists $h_0 > 0$ with

$$h_1 \equiv 1 - \eta \geq h_0 > 0, \quad h_2 \equiv \frac{1}{\delta} + \eta \geq h_0 > 0,$$
$$\gamma + \delta - \gamma \frac{(1+\delta^{-1})^2}{(h_1 + \gamma h_2)^3} |v|^2 \geq h_0 > 0. \tag{2.1}$$

There exists $T_{\max} > 0$ and a unique $U_{\rm RL} = (\eta, v)^\top \in C([0, T_{\max}); H^s(\mathbb{R})^2) \cap C^1([0, T_{\max}); H^{s-1}(\mathbb{R})^2)$, maximal solution to (1.2) (with $\epsilon = 1$), with initial data $U_{\rm RL}\,|_{t=0} = U^0$.
Moreover, there exists constants $0 < C_0, T^{-1} \leq \big|U^0\big|_{H^s(\mathbb{R})^2} C(\big|U^0\big|_{H^s(\mathbb{R})^2}, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ such that one has $T_{\max} \geq T$, and for any $t \in [0, T]$,

$$\big|U_{\rm RL}(t,\cdot)\big|_{H^s(\mathbb{R})^2} + \big|\partial_t U_{\rm RL}(t,\cdot)\big|_{H^{s-1}(\mathbb{R})^2} \leq C_0 \exp(C_0\, t),$$

and $U(t,\cdot)$ satisfies (2.1) uniformly for any $t \in [0, T]$ (with $h_0/2$ replacing h_0).

This result was precisely expressed in Guyenne et al. (2010), Theorem 1, and follows from standard techniques on quasilinear, Friedrichs-symmetrizable systems. More precisely, the existence and uniqueness of a solution follow from energy estimates on the linearized equation, of which the preceding estimate is a particular case. To assert the well-posedness in the sense of Hadamard, one should also state that the flow depends continuously upon the initial data. Such a result holds: one may control the energy of the difference between two solutions corresponding to different initial data, provided these initial data are sufficiently regular. Precise blow-up conditions, specifying the possible scenarios within the ones stated subsequently in Proposition 2.2 are also presented in Guyenne et al. (2010), Corollary 1.

Let us now turn to the free-surface system, (1.1). Recall that we set $\alpha = \varrho = \sqrt{\frac{1-\gamma}{\gamma+\delta}}$ and $\epsilon = 1$, so that the system may be written as

$$\partial_t U + A[U]\partial_x U = 0,$$

with $U \equiv (\zeta_1, \zeta_2, u_1, u_2)^\top$ and

$$A[U] \equiv \begin{pmatrix} u_1 & \frac{u_2 - u_1}{\varrho} & \frac{1+\varrho\zeta_1-\zeta_2}{\varrho} & \frac{\delta^{-1}+\zeta_2}{\varrho} \\ 0 & u_2 & 0 & \delta^{-1} + \zeta_2 \\ \frac{1}{\varrho} & 0 & u_1 & 0 \\ \frac{\gamma}{\varrho} & \delta + \gamma & 0 & u_2 \end{pmatrix} = A_0 + A_1(U),$$

where A_0 is a constant 4×4 matrix, and $A_1(\cdot)$ is a linear mapping into 4×4 matrices.

As we show in the appendix, the preceding system admits an explicit symmetrizer, $S[U]$, that is definite positive provided $U \equiv (\zeta_1, \zeta_2, u_1, u_2)^\top$ satisfies some conditions similar to (2.1), namely

$$\forall x \in \mathbb{R}, \qquad h_1(x) \ \geq \ h_0 \ > \ 0 \quad ; \quad h_2(x) - \frac{|u_2(x) - u_1(x)|^2}{\gamma + \delta} \ \geq \ h_0 \ > \ 0, \qquad (2.2)$$

where we recall $h_1 \equiv 1 + \varrho\zeta_1 - \zeta_2$ and $h_2 \equiv \delta^{-1} - \zeta_2$.

However, one clearly sees that the system exhibits $1/\varrho$ factors, which pass on the constants in the energy estimates, thereby lowering the a priori time of existence. We state subsequently in Proposition 2.2 the well-posedness of the Cauchy problem as given by standard energy methods on quasilinear, Friedrichs-symmetrizable systems; note that the time of existence of the solution is restricted to the poor $T_{\max} \gtrsim \varrho$. This time scale is intuitively seen from a change of variable: define $U(t, \cdot) \equiv \tilde{U}(t/\varrho, \cdot)$, so that $\tilde{U}$ satisfies

$$\partial_\tau \tilde{U} + \varrho A[\tilde{U}]\partial_x \tilde{U} \ = \ 0,$$

and one has $\varrho A[U] \equiv \varrho A_0 + \varrho A_1(U)$, with the matrix ϱA_0 and the linear mapping $\varrho A_1(\cdot)$ being both uniformly bounded with respect to $\varrho \ll 1$.

Proposition 2.2 *(Naive well-posedness result for the free-surface system) Let* $s \geq s_0 + 1$, $s_0 > 1/2$, *and let* $U^0 \equiv (\zeta_1^0, \zeta_2^0, u_1^0, u_2^0)^\top \in X^s$ *be such that* (2.2) *holds with* $h_0 > 0$.

There exist $T_{\max} > 0$ *and* $U = (\zeta_1^0, \zeta_2^0, u_1^0, u_2^0)^\top \in C([0, T_{\max}); H^s(\mathbb{R})^4) \cap C^1([0, T_{\max}); H^{s-1}(\mathbb{R})^4)$, *unique maximal solution to* (1.1) *(with* $\alpha = \varrho, \epsilon = 1$*), with initial data* $U\,|_{t=0} = U^0$.

Moreover, there exist positive constants $0 < C_0, T^{-1} \leq \big|U^0\big|_{X^s} C(\big|U^0\big|_{X^s}, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ *such that one has* $T_{\max} \geq T\varrho$, $U(t, \cdot)$ *satisfies* (2.2) *for any* $t \in [0, T\varrho]$ *(with* $h_0/2$ *replacing* h_0*), and*

$$\forall t \in [0, T\varrho], \qquad \big|U(t, \cdot)\big|_{X^s} + \varrho\big|\partial_t U(t, \cdot)\big|_{X^{s-1}} \ \leq \ C_0 \ \exp(C_0\, \varrho^{-1}\, t).$$

Finally, if $T_{\max} < \infty$*, then at least one of the following holds:*

- $\big|U\big|_{L^\infty([0,t]\times\mathbb{R})^4}$ *or* $\big|\partial_x U\big|_{L^\infty([0,t]\times\mathbb{R})^4}$ *blows up as* $t \nearrow T_{\max}$*; or*
- *at least one of the conditions in* (2.2) *ceases to be true at* $t = T_{\max}$.

The proof of Proposition 2.2 is postponed to the appendix so as not to interrupt the flow of the text.

Remark 2.3 Condition (2.2) is a sufficient condition for hyperbolicity, in the sense that it ensures that the symmetrizer we define and use in the appendix is positive definite. We do not claim that this condition defines exactly the domain of hyperbolicity of system (1.1) (contrary to (2.1) for the rigid-lid system (1.2)); see (Abgrall and Karni 2009; Castro-Díaz et al. 2011; Stewart and Dellar 2013) for a more detailed analysis on this point. In particular, one would expect the hyperbolic domain of the free-surface system to asymptotically correspond to (2.1) in the limit $\varrho \to 0$, which is not the case as (2.2) is more stringent.

Remark 2.4 Note that a uniform time of existence, $T \gtrsim 1$, is recovered for sufficiently small initial data: $\big|U_0\big|_{X^s} = \mathcal{O}(\varrho)$. This result can be viewed through the following change of unknowns: $U \equiv \varrho \breve{U}$. The function $\breve{U}$ satisfies

$$\partial_t \breve{U} + A[\varrho \breve{U}]\partial_x \breve{U} \;=\; 0,$$

and $A[\varrho \breve{U}] \equiv A_0 + \varrho A_1(\breve{U})$. The fact that the constant operator $A_0\partial_x$ is not uniformly bounded with respect to $\varrho \ll 1$ does not prevent solutions from existing in a time domain independent of ϱ because it does not contribute to commutator estimates. This simple observation motivates the strategy we use to prove Theorem 1.2, as described in Sect. 3.

3 Proof of Main Result

This section is dedicated to the proof of Theorem 1.2. Our first ingredient consists in constructing a system equivalent to (1.1) but whose nonlinear contribution is uniformly bounded with respect to ϱ. To construct such a system, we shall use different variables. Considering the conservation of horizontal momentum presented in Sect. 1.2, we introduce the horizontal momentum, $m \equiv \gamma h_1 u_1 + h_2 u_2$, and the shear velocity, $u_s \equiv u_2 - \gamma u_1$. One has immediately

$$u_s \equiv u_2 - \gamma u_1 \quad \text{and} \quad m \equiv \gamma h_1 u_1 + h_2 u_2 \tag{3.1}$$

if and only if

$$u_1 = \frac{m - h_2 u_s}{\gamma(h_1 + h_2)} \quad \text{and} \quad u_2 = \frac{m + h_1 u_s}{h_1 + h_2}. \tag{3.2}$$

Straightforward manipulations of system (1.1) yield the new system of conservation laws we consider:

$$\begin{cases} \partial_t \zeta_1 + \frac{1}{\varrho}\partial_x m + \frac{1-\gamma}{\gamma\varrho}\partial_x\left(h_1\frac{m-h_2u_s}{h_1+h_2}\right) = 0, \\ \partial_t\zeta_2 + \partial_x\left(\frac{h_2}{h_1+h_2}(h_1u_s + m)\right) = 0, \\ \partial_t u_s + (\delta+\gamma)\partial_x\zeta_2 + \dfrac{1}{2}\partial_x\left(\frac{\gamma(m+h_1u_s)^2-(m-h_2u_s)^2}{\gamma(h_1+h_2)^2}\right) = 0, \\ \partial_t m + \gamma\frac{h_1+h_2}{\varrho}\partial_x\zeta_1 + (\gamma+\delta)h_2\partial_x\zeta_2 + \partial_x\left(\frac{h_1(m-h_2u_s)^2+\gamma h_2(m+h_1u_s)^2}{\gamma(h_1+h_2)^2}\right) = 0. \end{cases} \tag{3.3}$$

We still refer to this system as a the *free-surface system*. Systems (3.3) and (1.1) are equivalent in the following sense.

Proposition 3.1 *Let* $s \geq s_0 + 1, s_0 > 1/2$. *Let* $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top \in C([0,T]; H^s(\mathbb{R})^4)$ *be a strong solution to* (3.3), *with* $T > 0$. *Assume that for any* $t \in [0,T]$ *one has*

$$\exists h_0 > 0 \text{ such that } \min_{x\in\mathbb{R}, t\in[0,T]} \left\{ h_1(t,x) + h_2(t,x) = 1 + \delta^{-1} + \varrho\zeta_1(t,x) \right\} \geq h_0 > 0.$$

Then $U \equiv (\zeta_1, \zeta_2, u_1, u_2)^\top \in C([0,T]; H^s(\mathbb{R})^4)$, *where* u_1 *and* u_2 *are given by* (3.2), *is a strong solution to* (1.1).

Conversely, if a given $U \equiv (\zeta_1, \zeta_2, u_1, u_2)^\top \in C([0,T]; H^s(\mathbb{R})^4)$ *is a strong solution to* (1.1) *and the preceding nonvanishing depth condition holds, then* $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top \in C([0,T]; H^s(\mathbb{R})^4))$, *given by* (3.1), *is a strong solution to* (3.3).

Proof The existence and regularity of $U \in C([0,T]; H^s(\mathbb{R})^4)$ (resp. $V \in C([0,T]; H^s(\mathbb{R})^4)$) are deduced from the corresponding control of V (resp. U), using product estimates in Lemma 5.1, as well as Corollary 5.2. As usual, one deduces from the system satisfied by, say, V—namely (3.3)—the corresponding estimate $\partial_t V \in C([0,T]; H^{s-1}(\mathbb{R})^4)$, and $\partial_t U \in C([0,T]; H^{s-1}(\mathbb{R})^4)$ follows. The fact that U satisfies (1.1) if V satisfies (3.3), and vice versa, demands somewhat tedious but straightforward computations, which we leave to the reader. □

Remark 3.2 We do not claim here that the aforementioned solutions are unique. The uniqueness of a solution to (1.1) is given in Proposition 2.2 and requires additional conditions on the initial data, namely (2.2). We prove subsequently that these conditions are also sufficient to ensure the uniqueness of a solution to (3.3); see Lemma 3.6.

Strategy and discussion. We see two benefits in considering (3.3) in lieu of (1.1). First, the rigid-lid system, which was encrypted in (1.1), is now apparent in (3.3). This will be helpful, although not necessary, for the construction of the approximate solution in the following subsection. More importantly, one sees that the only terms factored by ϱ^{-1} in (3.3) are constant. This second property is crucial for our analysis and justifies the use of (3.3).

Let us briefly sketch the key arguments in the proof of Theorem 1.2 before continuing with the detailed analysis in the following subsections. We first introduce some notations, which are used henceforth. We rewrite the hyperbolic system (3.3) as

$$\partial_t V + \left(\frac{1}{\varrho} L_\varrho + B[V] \right) \partial_x V = 0, \tag{3.4}$$

with $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top$, and where

- $\frac{1}{\varrho} L_\varrho$ represents the linear component of the system (see precise expression subsequently);
- $B[\cdot]$ contains a nonlinear contribution: it is uniformly bounded with respect to ϱ.

In Sect. 3.1, we construct an approximate solution, V_{app}, satisfying (3.3), as well as the initial data, up to a small remainder. Thus defining $W \equiv V - V_{\text{app}}$, where V is the exact solution, one has

$$\partial_t W \;+\; \frac{1}{\varrho}\big(L_\varrho + \varrho B[V_{\rm app} + W]\big)\,\partial_x W \;=\; \mathcal{R}, \tag{3.5}$$

with $W\,|_{t=0}$ and $\mathcal{R}$ small [typically of size $\mathcal{O}(\varrho)$]. Our aim is to prove that W remains small for large time scales (i.e., bounded from below uniformly with respect to ϱ), and Theorem 1.2 quickly follows (Sect. 3.3).

When compared with the classical theory of Friedrichs-symmetrizable quasilinear systems, the main issue we face when controlling W in the natural energy space lies in the two following facts:

(i) One must control the contribution from the unbounded component $\frac{1}{\varrho}L_\varrho$ in the energy space, which may generate a destructive $\mathcal{O}(\varrho^{-1})$ factor; and
(ii) One cannot use the equation to deduce a uniform control of $\partial_t W$ from the corresponding control of $\partial_x W$ because, once again, this would yield a destructive $\mathcal{O}(\varrho^{-1})$ factor.

These two difficulties are only apparent, as shown by a careful study of the system symmetrizer. In Sect. 3.2, we introduce and study the symmetrizer, $T[\cdot]$, as well as $\Upsilon[\cdot] \equiv T[\cdot]\big(\frac{1}{\varrho}L_\varrho + B[\cdot]\big)$. In particular, one can check that (roughly speaking) $\Upsilon[\cdot] \equiv \frac{1}{\varrho}\Upsilon_0 + \mathcal{O}(1)$, with Υ_0 a constant matrix, so that differentiation or commutation with the operator $\Upsilon[\cdot]$ is actually bounded; thus issue (i) can be addressed.

Issue (ii) requires a more specific analysis. We introduce $\Pi \equiv \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}$, the orthogonal projector onto the kernel of $L_{(0)}$, denoting $L_{(0)} = \lim_{\gamma\to 1} L_\varrho$; see below. It follows that $\big|\Pi\partial_t W\big|_{X^{s-1}} \lesssim \big|W\big|_{X^s}$, uniformly with respect to small ϱ. As for the other component, one shows that $T[\cdot](\mathrm{Id}-\Pi) = T_0 + \mathcal{O}(\varrho)$, with T_0 a constant matrix, so a factor of size $\mathcal{O}(\varrho)$ is gained following differentiation or commutation with this operator.

The detailed energy estimates are computed in Sect. 3.3.

There is an intuitive explanation for why the preceding claims hold. By precisely analyzing the 4×4 matrix L_ϱ,

$$L_\varrho \;\equiv\; \begin{pmatrix} 0 & 0 & \frac{\gamma-1}{\gamma(\delta+1)} & \frac{\gamma+\delta}{\gamma(\delta+1)} \\ 0 & 0 & \frac{\varrho}{1+\delta} & \frac{\varrho}{1+\delta} \\ 0 & \varrho(\gamma+\delta) & 0 & 0 \\ \gamma(1+\delta^{-1}) & \varrho\frac{\delta+\gamma}{\delta} & 0 & 0 \end{pmatrix},$$

one may check that for ϱ sufficiently small, L_ϱ has four distinct, real eigenvalues:

$$\lambda^f_\pm(\varrho) \;=\; \pm\sqrt{1+\delta^{-1}} \;+\; \mathcal{O}(\varrho^2) \quad;\quad \lambda^s_\pm(\varrho) \;=\; \pm\varrho \;+\; \mathcal{O}(\varrho^3).$$

The linear theory thus predicts that the flow can be decomposed as the superposition of four waves, propagating at velocity $c^f_\pm \sim \pm\frac{\sqrt{1+\delta^{-1}}}{\varrho}$, and $c^s_\pm \sim \pm 1$, which we call the *fast mode* (resp. *slow mode*). Roughly speaking, the slow mode corresponds to the flow predicted by the rigid-lid system, and the terms neglected in the rigid-lid approximation correspond to the fast mode.

An important feature of the free-surface system, which is revealed by our change of variable, is that the fast and slow modes are supported on (approximately) orthogonal components, which is responsible for the fact that coupling effects between the two modes are small. More precisely, if we use the notation $L_{(0)} \equiv \lim_{\gamma\to 1} L_\varrho \equiv \begin{pmatrix} & & & 1 \\ & & 0 & \\ & 0 & & \\ 1+\delta^{-1} & & & \end{pmatrix}$, then we can easily check that the eigenvectors corresponding to the two nonzero eigenvalues of $L_{(0)}$ are orthogonal to the kernel of $L_{(0)}$. Therefore, roughly speaking, the slow mode is supported by the variables ζ_2 and u_s, while the fast mode is supported by the variables ζ_1 and m. We take advantage of this fact by treating separately the slow-mode terms [multiplying by Π, the orthogonal projector onto the kernel of $L_{(0)}$] and fast-mode terms [multiplying by $\mathrm{Id}-\Pi$, the orthogonal projector onto the space spanned by the other eigenvectors of $L_{(0)}$]. The former contributions are easily controlled because time differentiation does not induce a destructive $\mathcal{O}(\varrho^{-1})$ factor. As for the latter, the property $T[\cdot](\mathrm{Id}-\Pi) = T_0 + \mathcal{O}(\varrho)$ reflects the fact that the corresponding eigenvalues are well separated; thus the perturbation by $\varrho B[\cdot]$ typically yields deviations of size $\mathcal{O}(\varrho)$, following standard perturbation theory Kato (1995). Finally, the desired property on $\Upsilon[\cdot]$ is easily checked:

$$\begin{aligned}\Upsilon[\cdot] &\equiv T[\cdot]\Pi\left(\frac{1}{\varrho}L_\varrho + B[\cdot]\right) + T[\cdot](\mathrm{Id}-\Pi)\left(\frac{1}{\varrho}L_\varrho + B[\cdot]\right)\\ &= \frac{1}{\varrho}T_0(\mathrm{Id}-\Pi)L_\varrho + \mathcal{O}(1).\end{aligned}$$

We refer the reader to Sect. 4 for a more precise investigation of the decomposition of the flow into fast and slow modes, and numerical illustrations.

3.1 Construction of Approximate Solution

In this section, we construct an approximate solution to the free-surface system (3.3) using the corresponding solution to the rigid-lid system (1.2), as defined below.

Let us recall that henceforth, we assume that γ is uniformly bounded from below: $\gamma \geq \gamma_{\min} > 0$. In particular, the norm X^s is equivalent to the standard $H^s(\mathbb{R})^4$-norm and will be used as such.

Definition 3.3 (Rigid-lid approximate solution) For given initial data ζ_2^0, u_s^0 satisfying (2.1), the *rigid-lid approximate solution* corresponding to $(\zeta_2^0, u_s^0)^\top$ is denoted $V_{\mathrm{RL}} \equiv (0, \eta, v, 0)^\top$, where $V \equiv (\eta, v)^\top$ is the unique solution to the rigid-lid system (1.2) with $V\mid_{t=0} \equiv (\zeta_2^0, u_s^0)^\top$.

Proposition 3.4 *Let $s \geq s_0$, $s_0 > 1/2$, and let $\zeta_2^0, u_s^0 \in H^{s+1}(\mathbb{R})$, satisfying (2.1) with $h_0 > 0$, and $\left|(\zeta_2^0, u_s^0)^\top\right|_{H^{s+1}\times H^{s+1}} \leq M$. Then there exists $0 < T^{-1}, C_1, C_2, C_3 \leq M\, C\big(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}\big)$, with the following statements:*

- $V_{\rm RL} \in C([0,T];X^{s+1}) \cap C^1([0,T];X^s)$ *is well defined as previously and satisfies*

$$\forall t \in [0,T], \qquad \big|V_{\rm RL}\big|_{X^{s+1}} + \big|\partial_t V_{\rm RL}\big|_{X^s} \ \le\ C_1. \tag{3.6}$$

- *There exists* $V_{\rm rem} \in C([0,T];X^{s+1}) \cap C^1([0,T];X^s)$, *with*

$$\forall t \in [0,T], \qquad \big|V_{\rm rem}\big|_{X^{s+1}} + \big|\partial_t V_{\rm rem}\big|_{X^s} \ \le\ C_2\,\varrho, \tag{3.7}$$

such that $V_{\rm app} \equiv V_{\rm RL} + V_{\rm rem}$ *satisfies* (3.3)*, up to a remainder term R, with*

$$\big\|R\big\|_{L^\infty([0,T];X^s)} \ \le\ C_3\,\varrho\,\big(M+\varrho\big). \tag{3.8}$$

Remark 3.5 The explicit formula for $V_{\rm rem}$, which is precisely displayed in the subsequent proof, does not play a significant role in this section, except as a technical artifice to obtain the desired estimate. In particular, it does not appear in Theorem 1.2. However, as discussed in Sect. 4, it corresponds to a first-order correction of the approximate solution and is clearly observable in our numerical simulations.

Proof of Proposition 3.4 By Proposition 2.1, there exists $C_1, T^{-1} \le MC(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ such that $V_{\rm RL} \in C([0,T];X^{s+1})$ is well defined by Definition 3.3, and (3.6) holds.

We now plug $V_{\rm app} \equiv V_{\rm RL} + V_{\rm rem}$ into (3.3) and check that one can explicitly define a function $V_{\rm rem} \equiv V_{\rm rem}[\eta, v]$ such that the remainder term, R, satisfies the estimate of the proposition. Anticipating the result, we use the notation $V_{\rm app} \equiv (\varrho\breve{\zeta}_1, \eta, v, \varrho^2\breve{m})^\top$ and subsequently

$$\begin{cases} \varrho\partial_t\breve{\zeta}_1 + \varrho\partial_x\breve{m} + \frac{1-\gamma}{\gamma\varrho}\partial_x\left(h_1\frac{\varrho^2\breve{m}-h_2 v}{h_1+h_2}\right) = r_1, \\ \partial_t\eta + \partial_x\left(\frac{h_2}{h_1+h_2}(h_1 v + \varrho^2\breve{m})\right) = r_2, \\ \partial_t v + (\delta+\gamma)\partial_x\eta + \frac{1}{2}\partial_x\left(\frac{\gamma(\varrho^2\breve{m}+h_1 v)^2 - (\varrho^2\breve{m}-h_2 v)^2}{\gamma(h_1+h_2)^2}\right) = r_3, \\ \varrho^2\partial_t\breve{m} + \gamma(h_1+h_2)\partial_x\breve{\zeta}_1 + (\gamma+\delta)h_2\partial_x\eta \\ +\partial_x\left(\frac{h_1(\varrho^2\breve{m}-h_2 v)^2 + \gamma h_2(\varrho^2\breve{m}+h_1 v)^2}{\gamma(h_1+h_2)^2}\right) = r_4, \end{cases} \tag{3.9}$$

with $h_1 \equiv 1 + \varrho^2\breve{\zeta}_1 - \eta$ and $h_2 \equiv \delta^{-1} + \eta$.

Our aim is to prove that one can choose $\breve{\zeta}_1$ and $\breve{m}$ such that

$$\big|\breve{\zeta}_1\big|_{H^{s+1}} + \big|\breve{m}\big|_{H^{s+1}} \ + \ \big|\partial_t\breve{\zeta}_1\big|_{H^s} + \big|\partial_t\breve{m}\big|_{H^s} \ \le\ C_2 \tag{3.10}$$

and

$$\big|r_1\big|_{H^s} + \big|r_2\big|_{H^s} + \big|r_3\big|_{H^s} + \big|r_4\big|_{H^s} \ \le\ C_3\,\varrho\,(M+\varrho). \tag{3.11}$$

To facilitate reading of the argument, we first assume that (3.10) holds and see how $\breve{\zeta}_1, \breve{m}$ can be naturally chosen so that (3.11) is satisfied. Our choice for $\breve{\zeta}_1, \breve{m}$ is precisely stated subsequently in (3.14) and (3.16), and checking that (3.10) is actually satisfied is then a straightforward consequence of (3.6).

Recall that, by definition, $(\eta, v)^\top$ satisfies (1.2). In particular, from the first equation in (1.2) one deduces

$$r_2 = \partial_x\left(\frac{h_1h_2v}{h_1+h_2} - \frac{\underline{h}_1h_2v}{\underline{h}_1+\gamma h_2}\right) + \varrho^2\partial_x\left(\frac{h_2\breve{m}}{h_1+h_2}\right),$$

where we denote by $\underline{h}_1 \equiv 1-\eta$ the depth of the upper layer in the rigid-lid approximation.

Recall that $V_{\rm RL}$ satisfies (3.6) and (2.1). Thus one can apply the product estimates in Lemma 5.1 and Corollary 5.2 [also recall that, by definition, $1-\gamma = \varrho^2(\gamma+\delta)$] to deduce

$$\big\|r_2\big\|_{L^\infty([0,T/M];H^s)} \le M\varrho^2\, C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}), \tag{3.12}$$

where we used the *a priori* estimate (3.10).

Similarly, one deduces from the second equation in (1.2) that

$$r_3 = \frac{1}{2}\partial_x\Bigg(\left\{\frac{\gamma h_1^2 - h_2^2}{\gamma(h_1+h_2)^2} - \frac{\underline{h}_1^2 - \gamma h_2^2}{(\underline{h}_1+\gamma h_2)^2}\right\}v^2 \\ + \frac{\gamma(\varrho^2\breve{m}+h_1v)^2 - \gamma^2h_1^2v^2 + h_2^2v^2 - (\varrho^2\breve{m}-h_2v)^2}{\gamma(h_1+h_2)^2}\Bigg),$$

so that one has, as previously,

$$\big\|r_3\big\|_{L^\infty([0,T/M];H^s)} \le M\varrho^2\, C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}). \tag{3.13}$$

Let us now look at the fourth equation in (3.9). We have

$$\gamma(h_1+h_2)\partial_x\breve{\zeta}_1 + (\gamma+\delta)h_2\partial_x\eta + \partial_x\left(\frac{h_1h_2(\gamma h_1+h_2)v^2}{\gamma(h_1+h_2)^2}\right) \\ = \partial_x\left(\gamma\big((1+\delta^{-1})\breve{\zeta}_1 + \frac{\varrho^2}{2}\breve{\zeta}_1^2\big) + (\gamma+\delta)\big(\delta^{-1}\eta + \frac{1}{2}\eta^2\big) + \frac{h_1h_2(\gamma h_1+h_2)v^2}{\gamma(h_1+h_2)^2}\right).$$

It is now clear that we can choose

$$\breve{\zeta}_1 \equiv -\big(\eta + \frac{\delta}{2}\eta^2\big) - \frac{(1-\eta)(\delta^{-1}+\eta)v^2}{(1+\delta^{-1})^2}, \tag{3.14}$$

so that the preceding expression is of size $\mathcal{O}(\varrho^2)$. More precisely, and using once again (3.10), we have

$$\big\|r_4\big\|_{L^\infty([0,T/M];H^s)} \le M\varrho^2\, C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}). \tag{3.15}$$

We conclude with the first equation in (3.9). Using that $\varrho^2 = \frac{1-\gamma}{\gamma+\delta}$, we have

$$r_1 = \varrho\left(\partial_t \breve{\zeta}_1 + \partial_x \breve{m} + \frac{\gamma+\delta}{\gamma}\partial_x\left(h_1 \frac{\varrho^2 \breve{m} - h_2 v}{h_1 + h_2}\right)\right).$$

We now recall that $(\eta, v)^\top$ satisfies (1.2), so that we deduce explicitly $\partial_t \breve{\zeta}_1$ from (3.14), and

$$\left|\partial_t \breve{\zeta}_1 - \partial_x\left(\frac{\underline{h}_1 h_2 v}{\underline{h}_1 + \gamma h_2}\right)\right|_{H^s} \le M^2\, C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}).$$

Now, we can check that by choosing

$$\breve{m} \equiv \frac{\delta}{1+\delta} v, \tag{3.16}$$

it follows that

$$\begin{aligned}&\left|\frac{\underline{h}_1 h_2 v}{\underline{h}_1 + \gamma h_2} + \breve{m} - \frac{\gamma+\delta}{\gamma}\frac{h_1 h_2 v}{h_1 + h_2}\right|_{H^s}\\ &\quad \le (M^2 + M\varrho^2)\, C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}),\end{aligned}$$

so that estimates (3.6) and (3.10) yield

$$\big\|r_1\big\|_{L^\infty([0,T/M];H^s)} \lesssim (M^2\varrho + M\varrho^2)\, C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}). \tag{3.17}$$

Estimates (3.12), (3.13), (3.15), and (3.17) give the desired estimate: (3.11) or, equivalently, (3.8). Moreover, one easily deduces from the estimate concerning V_{RL} in (3.6) the corresponding estimate on $V_{\mathrm{rem}} \equiv (\varrho\breve{\zeta}_1, 0, 0, \varrho^2\breve{m})^\top$: (3.10) or, equivalently, (3.7). Proposition 3.4 is proved. □

3.2 Properties of System and Its Symmetrizer

This section is dedicated to the preliminary results on the new free-surface system (3.3) and its symmetrizer, which will allow for an energy analysis in the following subsection.

Recall that (3.3) was constructed from (1.1) through a change of variables: for any $U \in (\zeta_1, \zeta_2, u_1, u_2)^\top$ solution to (1.1), we uniquely associate the $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top$ solution to (3.3) through a change of variable (3.1) (Lemma 3.1). In other words, we have an explicit

$$F : \begin{array}{rl} X & \to X \\ (\zeta_1, \zeta_2, u_s, m)^\top & \mapsto (\zeta_1, \zeta_2, u_1, u_2)^\top \end{array}$$

[in this section, the space X may be $L^\infty(\mathbb{R})^4$ or $H^s(\mathbb{R})^4$, $s > 1/2$], which is one-to-one and onto, provided the nonvanishing depth condition is satisfied:

$$\exists h_0 > 0 \ \ S.t. \min_{x\in\mathbb{R}, t\in[0,T]} \left\{h_1(t,x)+h_2(t,x)=1+\delta^{-1}+\varrho\zeta_1(t,x)\right\} \geq h_0 > 0. \quad (3.18)$$

It follows that, recalling the notation for (1.1) as

$$\partial_t U + A[U]\partial_x U = 0,$$

we may rewrite (3.3) (after multiplication by the appropriate operator)

$$\mathrm{d}F[V]\partial_t V + A[F(V)]\mathrm{d}F[V]\partial_x V = 0,$$

where $\mathrm{d}F[V]$ is the Jacobian matrix of F. In other words, recalling the earlier notation in (3.4), we have

$$\partial_t V + \left(\frac{1}{\varrho}L_\varrho + B[V]\right)\partial_x V = 0 \quad \text{with} \quad \frac{1}{\varrho}L_\varrho + B[V] = (\mathrm{d}F[V])^{-1}A[F(V)].$$

Thus the symmetrizer of the new system (3.3) is readily available from that of system (1.1).

Lemma 3.6 *Let $S[\cdot]$ be a symmetrizer of (1.1), e.g., (5.2). Then the operator $T[\cdot] \equiv (\mathrm{d}F[\cdot])^\top S[F(\cdot)]\mathrm{d}F[\cdot]$ is a symmetrizer of (3.3). Moreover, $T[V]$ is definite positive if and only if $F(V)$ satisfies (2.2).*

Proof For any $V \in X$, the operator $T[V]$ is obviously symmetric. Moreover, $T[V]$ is definite positive if and only if $S[F(V)]$ is definite positive since one has

$$\forall \mathbf{x} \in \mathbb{R}^4, \quad T[V]\mathbf{x}\cdot\mathbf{x} = S[F(V)](\mathrm{d}F[V]\mathbf{x})\cdot(\mathrm{d}F[V]\mathbf{x}), \quad (3.19)$$

and $\mathrm{d}F[V]$ is invertible provided V satisfies (3.18). Let us note that the hyperbolicity condition (2.2) is obviously more stringent than (3.18).

Finally, it is straightforward to check that

$$T[V]\left(\frac{1}{\varrho}L_\varrho + B[V]\right) = (\mathrm{d}F[V])^\top S[F(V)]A[F(V)]\mathrm{d}F[V]$$

is symmetric, and this concludes the proof. □

We conclude that one can construct an explicit symmetrizer of system (3.3) using $S[\cdot]$ given in (5.2). However, this symmetrizer has a quite complicated expression, and we do not display it here. We will only present the necessary properties of the operators of interest, which are easily checked using a computer algebra system, such as Maple.

Lemma 3.7 *Let $V, W \in X$, satisfying (3.18), and $\frac{1}{\varrho} L_\varrho + B[\cdot] \equiv (\mathrm{d}F[\cdot])^{-1} A[F(\cdot)] \mathrm{d}F[\cdot]$ defined previously. Then one has*

$$\big\| B[V] \big\|_X \leq C_0 \big| V \big|_X, \quad \big\| B[V] - B[W] \big\|_X \leq C_0 \big| V - W \big|_X, \tag{3.20}$$

with $C_0 = C(\big|V\big|_X, \big|W\big|_X, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$, and where we denote $\big\|A\big\|_X \equiv \sup_{V \in X \setminus \{\mathbf{0}\}} \frac{|AV|_X}{|V|_X}$.

Proof Recall that $B[\cdot]$ has a complicated expression, but it is explicit; it involves only products of the components of V or factors of the form $\frac{1}{h_1+h_2}$. Thus one can apply Lemma 5.1 and Corollary 5.2 [since (3.18) holds], and the result easily follows. □

Lemma 3.8 *Denote $T[\cdot] \equiv (\mathrm{d}F[\cdot])^\top S[F(\cdot)] \mathrm{d}F[\cdot]$ and $\Upsilon[\cdot] \equiv (\mathrm{d}F[\cdot])^\top \Sigma[F(\cdot)] \mathrm{d}F[\cdot]$, with $S[\cdot]$ and $\Sigma[\cdot] = S[\cdot]A[\cdot]$ defined in (5.2) and (5.3). Let $V \in X$ such that $F(V)$ satisfies (2.2), with $h_0 > 0$. Then there exists $C_0 = C(\big|V\big|_X, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$ such that one has the following claims:*

1. *$T[V], \Upsilon[V]$ are symmetric. $T[V]$ is positive definite. More precisely, for any $W \in L^2(\mathbb{R})^4$ one has*

$$\frac{1}{C_0} \big| W \big|_{L^2}^2 \leq \big(T[V]W, W \big) \leq C_0 \big| W \big|_{L^2}^2. \tag{3.21}$$

2. *$T[V], \Upsilon[V]$ satisfy the following estimates:*

$$\big\| T[V] \big\|_X \leq C_0; \quad \big\| \Upsilon[V] \big\|_X \leq \varrho^{-1} C_0. \tag{3.22}$$

3. *If $V \equiv V(\varkappa)$ and $\partial_\varkappa V \in X$, then*

$$\big\| \partial_\varkappa (T[V]) \big\|_X \leq C_0 \big| \partial_\varkappa V \big|_X; \quad \big\| \partial_\varkappa (\Upsilon[V]) \big\|_X \leq C_0 \big| \partial_\varkappa V \big|_X \tag{3.23}$$

and

$$\big\| \partial_\varkappa (T[V])(\mathrm{Id} - \Pi) \big\|_X \leq \varrho\, C_0 \big| \partial_\varkappa V \big|_X, \tag{3.24}$$

recalling the notation $\Pi \equiv \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}$.

Proof That $T[V]$ is symmetric, positive definite and $\Upsilon[V]$ is symmetric was already stated in Lemma 3.6. Estimate (3.21) follows from Lemma 5.4 and (3.19), recalling that $\gamma \geq \gamma_{\min} > 0$ ensures that the $L^2(\mathbb{R})^4$-norm is equivalent to the X^0-norm.

Estimates (3.22) are direct consequences of the corresponding estimates on $S[\cdot]$, $A[\cdot]$ and $F(\cdot)$, $\mathrm{d}F[\cdot]$, which are easily checked. We recall that the necessary product estimates in $X = L^\infty(\mathbb{R})^4$ or $X = H^s(\mathbb{R})^4$ ($s > 1/2$) are given by Lemma 5.1 and Corollary 5.2. The first estimate in (3.23) is obtained similarly.

Finally, the second estimate in (3.23) and (3.24) are less obvious but can be checked with the help of a computer algebra system (we must ensure that all first-order terms in ϱ are constant). □

3.3 Completion of Proof

Denote by V a strong solution to the free-surface system (3.3) satisfying the nonvanishing depth condition, (2.1); and denote by $V_{\rm app}$ the approximate solution constructed in Proposition 3.4. One easily checks that $W \equiv V - V_{\rm app}$ satisfies the following system:

$$\partial_t W \ + \ \frac{1}{\varrho}\left(L_\varrho + \varrho B[V]\right)\partial_x W \ = \ \mathcal{R}, \tag{3.25}$$

with $\mathcal{R} \equiv R-(B[V_{\rm app}+W]-B[V_{\rm app}])\partial_x V_{\rm app}$, where R is estimated in Proposition 3.4.

The following lemma presents an a priori energy estimate on W satisfying the preceding system, on which our desired result is based.

Lemma 3.9 *Let $s \geq s_0 + 1,\ s_0 > 1/2$, and let W be a strong solution to (3.25), with $W\,|_{t=0} \in X^s$. Assume that there exists $M, T, h_0 > 0$ such that $F(V)$ satisfies (2.2) and*

$$\big\|V\big\|_{L^\infty([0,T];X^s)} \ + \ \big\|\partial_t V\big\|_{L^\infty([0,T];X^{s-1})} \ \leq \ M.$$

Then one has, for all $t \in [0, T]$,

$$\big|W(t,\cdot)\big|_{X^s} \leq C_0\big|W(0,\cdot)\big|_{X^s} e^{C_0 M t} \ + \ C_0 \int_0^t e^{C_0 M(t-t')}\big|\mathcal{R}(t',\cdot)\big|_{X^s}\, dt', \tag{3.26}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Proof We compute the inner product of (3.25) with $T[V]\Lambda^{2s}W$ and obtain

$$\left(\Lambda^s T[V]\partial_t W, \Lambda^s W\right) \ + \ \left(\Lambda^s \Upsilon[V]\partial_x W, \Lambda^s W\right) \ = \ \left(\Lambda^s T[V]\mathcal{R}, \Lambda^s W\right),$$

where $T[\cdot]$ and $\Upsilon[\cdot]$ are as defined in the previous subsection.

From the symmetry of $T[\cdot]$ and $\Upsilon[\cdot]$ one deduces

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}E^s(W) \ &= \frac{1}{2}\left(\left[\partial_t, T[V]\right]\Lambda^s W, \Lambda^s W\right) + \frac{1}{2}\left(\left[\partial_x, \Upsilon[V]\right]\Lambda^s W, \Lambda^s W\right) \\ &- \left(\left[\Lambda^s, T[V]\right]\partial_t W, \Lambda^s W\right) \ - \ \left(\left[\Lambda^s, \Upsilon[V]\right]\partial_x W, \Lambda^s W\right) + \left(\Lambda^s T[V]\mathcal{R}, \Lambda^s W\right),\end{aligned} \tag{3.27}$$

where we define

$$E^s(W) \equiv \left(T[V]\Lambda^s W, \Lambda^s W\right).$$

We estimate below each of the terms on the right-hand side of (3.27).

Estimate of $\big([\partial_t, T[V]]\Lambda^s W, \Lambda^s W\big)$. From (3.23) in Lemma 3.8 [with $X = L^\infty(\mathbb{R})^4$] we have

$$\big|[\partial_t, T[V]]\Lambda^s W\big|_{L^2} \le \big|\partial_t V\big|_{L^\infty} C\big(|V|_{L^\infty}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}\big)\big|\Lambda^s W\big|_{L^2}.$$

By hypothesis, $\big|\partial_t V\big|_{X^{s-1}}$ is controlled, and continuous Sobolev embeddings for $s - 1 \ge s_0 > 1/2$ imply an equivalent control on the L^∞-norm. One obtains simply

$$\big|[\partial_t, T[V]]\Lambda^s W\big|_{L^2} \le M\, C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})\, \big|\Lambda^s W\big|_{L^2}.$$

It follows from the preceding expression and the Cauchy–Schwarz inequality that

$$\big|\big([\partial_t, T[V]]\Lambda^s W, \Lambda^s W\big)\big| \le C_0\, M\, \big|W\big|_{X^s}^2, \tag{3.28}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Estimate of $\big([\partial_x, \Upsilon[V]]\Lambda^s W, \Lambda^s W\big)$. As previously, the Cauchy–Schwarz inequality and Lemma 3.8 yield

$$\big([\partial_x, \Upsilon[V]]\Lambda^s W, \Lambda^s W\big) \le \big|\partial_x V\big|_{L^\infty} C\big(|V|_{L^\infty}, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}\big)\big|\Lambda^s W\big|_{L^2}^2,$$

which is easily estimated thanks to continuous Sobolev embeddings. One obtains

$$\big|\big([\partial_x, \Upsilon[V]]\Lambda^s W, \Lambda^s W\big)\big| \le C_0\, M\, \big|W\big|_{X^s}^2, \tag{3.29}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Estimate of $\big(\Lambda^s T[V]\mathcal{R}, \Lambda^s W\big)$. We apply the Cauchy–Schwarz inequality and (3.22) in Lemma 3.8. One deduces

$$\big(\Lambda^s T[V]\mathcal{R}, \Lambda^s W\big) \le C_0\big|W\big|_{X^s}\big|\mathcal{R}\big|_{X^s}, \tag{3.30}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Estimate of $\big([\Lambda^s, \Upsilon[V]]\partial_x W, \Lambda^s W\big)$. We make use of the Kato–Ponce commutator estimate recalled in Lemma 5.3. It follows that

$$\big|[\Lambda^s, \Upsilon[V]]\partial_x W\big|_{L^2(\mathbb{R})^4} \lesssim \big\|\partial_x(\Upsilon[V])\big\|_{X^{s-1}}\big|\partial_x W\big|_{X^{s-1}}.$$

From (3.23) in Lemma 3.8, and since X^{s-1} is a Banach algebra, one has

$$\begin{aligned}\big\|\partial_x(\Upsilon[V])\big\|_{X^{s-1}} &\lesssim \big|\partial_x V\big|_{X^{s-1}} C\big(|V|_{X^{s-1}}, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}\big)\\ &\lesssim M\, C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}).\end{aligned}$$

It follows that

$$\big|\big([\Lambda^s, \Upsilon[V]]\partial_x W, \Lambda^s W\big)\big| \le C_0\, M\, \big|W\big|_{X^s}^2, \tag{3.31}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Estimate of $\big(\big[\Lambda^s, T[V]\big]\partial_t W, \Lambda^s W\big)$. As previously, the Kato–Ponce commutator estimate yields

$$\big|\big[\Lambda^s, T[V]\big]\partial_t W\big|_{L^2(\mathbb{R})^4} \lesssim \big\|\partial_x(T[V])\big\|_{X^{s-1}}\big|\partial_t W\big|_{X^{s-1}} \lesssim M\big|\partial_t W\big|_{X^{s-1}}.$$

Unfortunately, making use of the identity (3.25) only yields $\big|\partial_t W\big|_{X^{s-1}} \lesssim \frac{1}{\varrho}\big|W\big|_{X^s}$, which is not sufficient to conclude. Thus we now need to use precisely the structure of our system, and in particular estimate (3.24). Thus we decompose $[\Lambda^s, T[V]]\partial_t W$ into two components:

$$\big[\Lambda^s, T[V]\big]\partial_t W \;\equiv\; \big[\Lambda^s, T[V]\big]\Pi\partial_t W + \big[\Lambda^s, T[V]\big](\mathrm{Id} - \Pi)\partial_t W.$$

Let us start with the "slow" contribution, $\big[\Lambda^s, T[V]\big]\Pi\partial_t W$. One can use Eq. (3.25) to control $\Pi\partial_t W$ uniformly with respect to small ϱ. Indeed, one has

$$\Pi\partial_t W \;=\; -\frac{1}{\varrho}\Pi L_\varrho \partial_x W - \Pi B[V]\partial_x W + \Pi\mathcal{R},$$

so that

$$\begin{aligned}\big|\Pi\partial_t W\big|_{X^{s-1}} \;&\leq\; \big|\frac{1}{\varrho}\Pi L_\varrho\partial_x W\big|_{X^{s-1}} \;+\; \big|B[V]\partial_x W\big|_{X^{s-1}} \;+\; \big|\mathcal{R}\big|_{X^{s-1}},\\ &\leq\; (1+M)C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})\big|\partial_x W\big|_{X^{s-1}} + \big|\mathcal{R}\big|_{X^{s-1}},\end{aligned}$$

where we used estimate (3.22) in Lemma 3.8 and the property $\big\|\Pi L_\varrho\big\| = \mathcal{O}(\varrho)$. It follows that

$$\big|\big[\Lambda^s, T[V]\big]\Pi\partial_t W\big|_{L^2(\mathbb{R})^4} \leq C_0\; M\;\big(\big|W\big|_{X^s} + \big|\mathcal{R}\big|_{X^{s-1}}\big),$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

We continue with the "fast" contribution, $\big[\Lambda^s, T[V]\big](\mathrm{Id} - \Pi)\partial_t W$. Since $(\mathrm{Id} - \Pi)$ is constant, it commutes with Λ^s, and Kato–Ponce commutator estimates (Lemma 5.3) yield

$$\big|\big[\Lambda^s, T[V]\big](\mathrm{Id} - \Pi)\partial_t W\big|_{L^2(\mathbb{R})^4} \lesssim \big\|\partial_x\,(T[V](\mathrm{Id} - \Pi))\,\big\|_{H^{s-1}}\big|\partial_t W\big|_{X^s}.$$

Now one has as previously

$$\big|\partial_t W\big|_{X^s} \leq C_0\big(\frac{1}{\varrho}\big|W\big|_{X^s} + \big|\mathcal{R}\big|_{X^{s-1}}\big),$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$. Estimate (3.24) in Lemma 3.8 makes it possible to recover a factor of size $\mathcal{O}(\varrho)$:

$$\big\|\partial_x\,(T[V](\mathrm{Id} - \Pi))\,\big\|_{H^{s-1}} \;\leq\; C_0\; M\;\varrho,$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$. Thus we proved

$$\big|\big[\Lambda^s, T[V]\big]\Pi\partial_t W\big|_{L^2(\mathbb{R})^4} \lesssim C_0\, M\, \big(|W|_{X^s} + \varrho|\mathcal{R}|_{X^{s-1}}\big),$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Altogether, one has, applying the Cauchy–Schwarz inequality,

$$\big|\big(\big[\Lambda^s, T[V]\big]\partial_t W, \Lambda^s W\big)\big| \;\leq\; C_0\, M\, \big(|W|_{X^s} + |\mathcal{R}|_{X^{s-1}}\big)|W|_{X^s}\;, \tag{3.32}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Plugging (3.28)–(3.32) into (3.27) yields

$$\frac{1}{2}\frac{d}{dt}E^s(W) \;\leq\; C_0\big(M\,|W|_{X^s}^2 \;+\; |\mathcal{R}|_{X^s}|W|_{X^s}\big).$$

Finally, estimate (3.21) in Lemma 3.8 yields

$$\frac{1}{2}\frac{d}{dt}E^s(W) \;\leq\; C_0'\, M\, E^s(W) \;+\; C_0'|\mathcal{R}|_{X^s}E^s(W)^{1/2},$$

with $C_0' = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$, and Lemma 3.9 follows from the Gronwall–Bihari lemma. □

Completion of proof of Theorem 1.2. Let us now quickly show how Theorem 1.2 follows from Lemma 3.9. For given initial data as in the theorem, Proposition 2.2 yields the existence of $T_{\max} > 0$ and a unique solution $U \equiv (\zeta_1, \zeta_2, u_1, u_2)^\top \in C([0, T_{\max}); X^{s+1}) \cap C^1([0, T_{\max}); X^s)$ to (1.1) such that $U(t, \cdot)$ satisfies (2.2) for $t \in [0, T_{\max})$. It follows from Proposition 3.1 that the change of variables in (3.1) yields $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top \in C([0, T_{\max}); X^{s+1}) \cap C^1([0, T_{\max}); X^s)$ as a solution to (3.3).

Thanks to Proposition 3.4, and since condition (1.6) ensures that $(\zeta_2^0, u_s^0)^\top$ satisfies (2.1), one has that $V_{\text{app}} = V_{\text{RL}} + V_{\text{rem}}$ is well defined and controlled for $t \in [0, T/M]$. More precisely, there exists $T^{-1}, C_1 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$ such that

$$\sup_{t\in[0,T/M]}\big\{\big|V_{\text{app}}(t,\cdot)\big|_{X^{s+1}} + \big|\partial_t V_{\text{app}}(t,\cdot)\big|_{X^s}\big\} \;\leq\; C_1 M. \tag{3.33}$$

Denote $W \equiv V - V_{\text{app}}$. By construction, one has

$$\big|W\,|_{t=0}\,\big|_{X^s} + \varrho\big|\partial_t W\,|_{t=0}\,\big|_{X^{s-1}} \;\leq\; C_2\,\varrho\, M,$$

with $C_2 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$. We introduce the time $T^\sharp$ as

$$T^\sharp \equiv \sup\Big\{t \in [0, T_{\max}, T/M],\ \|W\|_{L^\infty([0,t];X^s)} + \varrho\|\partial_t W\|_{L^\infty([0,t];X^{s-1})} \leq 2C_2\,\varrho\, M\Big\}. \tag{3.34}$$

One has $T^\sharp > 0$ since $W = V - V_{\rm app} \in C([0,T]; X^{s+1}) \cap C^1([0,T]; X^s)$; our aim is to prove that $T^\sharp$ is uniformly bounded from below as in Theorem 1.2.

Recall that W satisfies (3.25); thus we apply Lemma 3.9 with

$$\mathcal{R} \equiv R - (B[V_{\rm app} + W] - B[V_{\rm app}])\partial_x V_{\rm app}.$$

Proposition 3.4 yields

$$\big\| R \big\|_{L^\infty([0,T/M];X^s)} \lesssim M\varrho\,\big(M + \varrho\big).$$

Now, using that $(X^s, \big|\cdot\big|_{X^s})$ is a Banach algebra, and using (3.20) in Lemma 3.7, one has

$$\big|(B[V_{\rm app} + W] - B[V_{\rm app}])\partial_x V_{\rm app}\big| \lesssim C(\big|V_{\rm app}\big|_{X^s}, \big|W\big|_{X^s})\big|W\big|_{X^s}\big|\partial_x V_{\rm app}\big|_{X^s}.$$

It follows from the preceding estimates that

$$\big\|\mathcal{R}\big\|_{L^\infty([0,T^\sharp];X^s)} \leq C_3(M^2\varrho + M\varrho^2), \tag{3.35}$$

with $C_3 = C(M, h_0^{-1}\delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Finally, we apply (3.26) in Lemma 3.9 [making use of (3.33)–(3.35)] and deduce

$$\forall\, 0 \leq t \leq T^\sharp, \qquad \big|W(t,\cdot)\big|_{X^s} \leq C_0 M\varrho e^{C_0 M t} + C_0(M\varrho + \varrho^2)(e^{C_0 M t} - 1)\ ,$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$. A similar estimate is obtained on $\partial_t W$ using the equation satisfied by W, namely (3.25):

$$\big|\partial_t W\big|_{X^{s-1}} \leq \frac{1}{\varrho} C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})\big|\partial_x W\big|_{X^{s-1}} + \big|\mathcal{R}\big|_{X^{s-1}}.$$

It follows that there exists $T' > 0$, depending nondecreasingly on $M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}$, such that one has

$$T^\sharp \geq \min\{T_{\max}, T'/M, T'/\varrho\}.$$

Triangular inequalities and (3.33), (3.34) immediately yield

$$\big\|\zeta_2\big\|_{L^\infty([0,T^\sharp];H^s)} + \big\|u_s\big\|_{L^\infty([0,T^\sharp];H^s)} \leq M\exp(C_0 M t), \tag{3.36}$$

$$\big\|\zeta_1\big\|_{L^\infty([0,T^\sharp];H^s)} + \big\|m\big\|_{L^\infty([0,T^\sharp];H^s)} \leq M\varrho\exp(C_0 M t), \tag{3.37}$$

$$\big\||\partial_t\zeta_1| + |\partial_t\zeta_2| + |\partial_t u_s| + |\partial_t m|\big\|_{L^\infty([0,T^\sharp];H^{s-1})} \leq M\exp(C_0 M t), \tag{3.38}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

It follows in particular from (3.38) that for any $t \in [0, T^\sharp]$ one has

$$\begin{aligned}\big|h_2(t,\cdot) - h_2(0,\cdot)\big|_{H^{s-1}} &\leq \Big| \int_0^t \partial_t \zeta_2(t',\cdot)\,dt' \Big|_{H^{s-1}} \\ &\leq C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})\, M\, t,\end{aligned}$$

where we recall that $h_2 \equiv \delta^{-1} + \zeta_2$. Similar estimates on $h_1 \equiv 1 + \varrho\zeta_1 - \zeta_2$ and u_1, u_2 given by (3.2) show that $U(t,\cdot)$ satisfies condition (2.2) uniformly for $t \in [0, \min\{T^\sharp, T''/M\})$ (replacing h_0 with $h_0/2$), with $T''^{-1} = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

From the blow-up conditions stated in Proposition 2.2 and a classical continuity argument, it is now clear that there exists $T > 0$, depending only and nondecreasingly on $M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}$ such that $T_{\max} \geq T/\max\{M, \varrho\}$.

The estimates in Theorem 1.2 are a straightforward consequence of (3.34), (3.36), and (3.37) (using Lemma 5.1 and Corollary 5.2), and the proof of Theorem 1.2 is now complete.

4 Decomposition of the Flow

In this section, we offer partial answers to two of the natural questions arising from Theorem 1.2:

1. *Can we describe more precisely the asymptotic behavior of the solution, and in particular the leading-order deformation of the surface?*
2. *Can we extend the result to ill-prepared initial data, that is, data that fail to meet the smallness assumption in* (1.5)?

In both cases, as we shall see, the answer will be given through a decomposition between fast and slow modes. Such a decomposition is exact in the linear case [$\epsilon = 0$ in (1.1)] as the system becomes a linear wave equation; therefore the flow is a superposition of four traveling waves. Diagonalizing $\frac{1}{\varrho}L_\varrho$ [using the notation introduced in (3.4)] shows that when $\varrho \to 0$, two of these waves (corresponding to the solution of the rigid-lid system and mainly supported on variables ζ_2 and u_s) traveling at a velocity $c_\pm^s \sim \pm 1$, while the two other ones (mainly supported on ζ_1 and m) travel at a velocity $c_\pm^f \sim \pm\sqrt{1+\delta^{-1}}/\varrho$.

This decomposition is hardly new. In the literature, the two modes are also often referred to as surface/interface modes, or barotropic/baroclinic modes, since the fast-mode components share the properties of water waves for one layer of a fluid of constant mass density (Gill 1982). The decomposition is exact in the linear setting and has been shown to hold approximately in the weakly nonlinear setting; see Duchêne (2011) and references therein. In that case, the smallness of ϵ makes it possible to control the coupling effects between each of the waves (even when additional—small—dispersion terms are included), provided the initial data are sufficiently spatially localized.

Our aim in this section is to show that this decomposition is quite robust and holds even when strong nonlinearities are involved. As was already mentioned, such a result will rely on a condition of spatial localization of the initial data, which we express through weighted Sobolev spaces.

In Sect. 4.1, we construct slow- and fast-mode correctors that make it possible to obtain a higher-order approximate solution of the free-surface system using only the corresponding solution to the rigid-lid system and the initial data. Thus we improve the results stated in Proposition 3.4 and Theorem 1.2 with Proposition 4.2 and Theorem 4.5, respectively. In Sect. 4.2, we extend the consistency result obtained in Proposition 3.4 to ill-prepared initial data, that is, data allowing nonsmall horizontal momentum and deformation of the surface and, thus, involving a leading-order slow mode. Unfortunately, we cannot conduct a study as in Sect. 3.3 and deduce the stronger result corresponding to Theorem 1.2 (although numerical simulations are in full agreement with such results). Finally, Sect. 4.3 contains numerical simulations illustrating the aforementioned results and an accompanying discussion.

Remark 4.1 Recall that we set $\epsilon = 1$ and $\alpha = \varrho$ after Theorem 1.2; see Remarks 1.3 and 1.4. The general setting and, therefore, statements as in Theorem 1.2 are easily recovered. We also implicitly assume that the constant M, which evaluates the magnitude of the initial perturbation, is bounded from below. More specifically, for technical reasons, we restrict our study to the time interval $t \in [0, T]$, with T^{-1} bounded, rather than $t \in [0, T/M]$—although, as discussed in Remark 1.3, we do not expect any particular limitation to occur when M is small.

As for Theorem 1.2 (Remark 1.5), our statements do not require the parameter ϱ to be small but are of little interest otherwise. In particular, our strategy of approximating the flow as the superposition of fast- and a slow-mode approximate solutions relies heavily on the fact that the fast mode propagates at a velocity $|c| \gtrsim 1/\varrho$, so that coupling effects are strong only during a time interval of size $\mathcal{O}(\varrho)$ (since the two modes are localized away from each other afterward).

If both M and ϱ are not small, then the initial perturbation will give rise to fast and slow modes of comparable magnitude and velocity. The two modes will therefore interact in a nontrivial, nonlinear way, and the full free-surface system is required to accurately describe the flow.

4.1 Improved Approximate Solution

In this section, we show that one can construct a first-order corrector to the rigid-lid approximate solution given in Theorem 1.2, provided the initial data are bounded in weighted Sobolev spaces. A key ingredient is the establishment of a fast-mode corrector, which makes it possible to take into account small initial data supported on the variables ζ_1 and m.

In Proposition 4.2, we provide a higher-order approximate solution to (3.3) in the sense of consistency, i.e., similarly to Proposition 3.4. One can then apply the strategy developed in Sect. 3, and one obtains the stronger result expressed subsequently in Theorem 4.5.

Proposition 4.2 *Let $s \geq s_0,\ s_0 > 1/2$, and let $\zeta_1^0, \zeta_2^0, u_s^0, m^0 \in H^{s+1}(\mathbb{R})$, satisfying (1.5),(1.6) [following the change of variable in (3.2)] with given $0 < M, h_0 < \infty$.*

Assume additionally that there exists $\sigma > 1/2$ *such that*

$$\big|(1+|\cdot|^2)^\sigma \zeta_1^0\big|_{H^{s+1}} + \big|(1+|\cdot|^2)^\sigma m^0\big|_{H^{s+1}} + \varrho\big|(1+|\cdot|^2)^\sigma \zeta_2^0\big|_{H^{s+1}} + \varrho\big|(1+|\cdot|^2)^\sigma u_s^0\big|_{H^{s+1}} \ \le\ M\varrho.$$

Then there exists $0 < T^{-1}, C_0 \le C(M, h_0^{-1}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$ *such that*

1. $V_{\mathrm{RL}} \equiv (0,\eta,v,0)^\top$ *is well defined by Definition 3.3 and satisfies*

$$\forall t \in [0,T], \qquad \big|V_{\mathrm{RL}}\big|_{X^{s+1}} + \big|\partial_t V_{\mathrm{RL}}\big|_{X^s} \ \le\ C_0\, M;$$

2. $V_{\mathrm{cor}}^s \equiv (\varrho\breve{\zeta}_1, 0, 0, 0)^\top$ *is well defined, with*

$$\breve{\zeta}_1 \ \equiv\ -\Big(\eta + \frac{\delta}{2}\eta^2\Big) - \frac{(1-\eta)(\delta^{-1}+\eta)v^2}{(1+\delta^{-1})^2};$$

3. V_{cor}^f *is well defined, with*

$$V_{\mathrm{cor}}^f(t,x) \equiv \begin{pmatrix} u_+(x-c/\varrho t) + u_-(x+c/\varrho t) \\ 0 \\ 0 \\ c(u_+(x-c/\varrho t) - u_-(x+c/\varrho t)) \end{pmatrix},$$

where $c \equiv \sqrt{1+\delta^{-1}}$ *and* $u_\pm(x) \ = \ \frac{1}{2}\big(\zeta_1^0 - \varrho\breve{\zeta}_1\,|_{t=0} \pm c^{-1}m^0\big)$*;*

4. *There exists* V_{rem}*, with*

$$\forall t \in [0,T], \qquad \big|V_{\mathrm{rem}}(t,\cdot)\big|_{X_{\mathrm{ul}}^{s+1}} \ \le\ C_0\, M\,,$$

such that $V_{\mathrm{app}} \equiv V_{\mathrm{RL}} + V_{\mathrm{cor}}^s + V_{\mathrm{cor}}^f + \varrho^2 V_{\mathrm{rem}}$ *satisfies (3.3) up to a remainder term,* R*, with*

$$\int_0^T \big|R(t,\cdot)\big|_{X^s}\, dt \ \le\ C_0\, M\, \varrho^2.$$

Remark 4.3 We denote by $(H_{\mathrm{ul}}^s, |\cdot|_{H_{\mathrm{ul}}^s})$ the uniformly local Sobolev space introduced in Kato (1975):

$$\big|u\big|_{H_{\mathrm{ul}}^s} \ \equiv\ \sup_{j\in\mathbb{N}} \big|\chi(\cdot - j)u(\cdot)\big|_{H^s},$$

where χ is a smooth function satisfying $\chi \equiv 0$ for $|x| \ge 1$, $\chi \equiv 1$ for $|x| \le 1/2$, and $\sum_{j\in\mathbb{N}} \chi(x-j) = 1$ for any $x \in \mathbb{R}$ (the space is independent of the choice of χ satisfying these assumptions).

We then use the notation $(X^s_{\rm ul}, \big|\cdot\big|_{X^s_{\rm ul}})$ and $(L^\infty([0,T];X^s_{\rm ul}), \big\|\cdot\big\|_{L^\infty([0,T];X^s_{\rm ul})})$, similarly to the previously defined Sobolev-based spaces.

Proof of Proposition 4.2 The well-posedness and estimate of $V_{\rm RL}$ for $t\in[0,T]$ was stated in Proposition 3.4 (here and in what follows, unless otherwise stated, we use the notation $T=\tilde{T}/M$, where $\tilde{T}$ is the constant used for the time intervals in the statements of Sect. 3). The definition of the corrector and remainder terms, as well as the desired estimates, is obtained in three steps. First, we construct a high-order approximate solution corresponding to the initial data ζ_2^0, u_s^0, using the corresponding solution to the rigid-lid system, that we will refer to as the *slow-mode approximate solution*. Then we see how to construct the *fast-mode approximate solution* in order to deal with the inadequacy of the slow-mode approximate solution with regard to the initial data. Finally, we show that, thanks to the localization in space of the initial data, the coupling effects between the two modes are weak, so that the superposition of the two contributions produces the desired approximate solution.

Construction of slow-mode approximate solution. We proceed as in the proof of Proposition 3.4, but we propose a higher-order definition for the corrector term in order to attain the improved precision. More precisely, we seek $V^s_{\rm app}\equiv V_{\rm RL}+V^s_{\rm cor}+\varrho^2 V_{\rm rem}$, with $V_{\rm RL}+V^s_{\rm cor}\equiv(\varrho\breve{\zeta}_1,\eta,v,0)^\top$ as in the proof of Proposition 3.4 and $V_{\rm rem}\equiv(0,0,0,\breve{m})^\top$ to be determined. Following the same steps as in the proof of Proposition 3.4, we see that the only difficulty we face lies in the estimate of

$$r_1 \ = \ \varrho\left(\partial_t\breve{\zeta}_1+\partial_x\breve{m}+\frac{\gamma+\delta}{\gamma}\partial_x\left(h_1\frac{\varrho^2\breve{m}-h_2 v}{h_1+h_2}\right)\right),$$

where $V_{\rm RL}\equiv(0,\eta,v,0)^\top$ is the rigid-lid solution defined in Definition 3.3 and $\breve{\zeta}_1$ is as defined in (3.14). It is therefore natural to set

$$\breve{m}(t,x) \ \equiv \ -\int_0^x \partial_t\breve{\zeta}_1(t,x')\,dx'+\delta\underline{h}_1(t,x)h_2(t,x)v(t,x), \tag{4.1}$$

where we use the notation $\underline{h}_1\equiv 1-\eta$ and $h_2\equiv\delta^{-1}+\eta$.

Note that $\breve{m}$ may not have finite energy since it does not necessarily decay when $x\to\pm\infty$. However, recall the estimates of Proposition 3.4:

$$\forall t\in[0,T],\qquad \big|V_{\rm RL}\big|_{X^{s+1}}+\big|\partial_t V_{\rm RL}\big|_{X^s}\ \lesssim\ C_0\,M, \tag{4.2}$$

$$\forall t\in[0,T],\qquad \big|\breve{\zeta}_1\big|_{H^{s+1}}+\big|\partial_t\breve{\zeta}_1\big|_{H^s}\ \lesssim\ C_0\,M. \tag{4.3}$$

[here and in what follows, we use the notation $C_0=C(M,h_0^{-1},\delta_{\min}^{-1},\delta_{\max},\gamma_{\min}^{-1})$]. One deduces

$$\forall t\in[0,T],\qquad \big|\breve{m}\big|_{H^{s+1}_{\rm ul}}\ +\ \big|\partial_x\breve{m}\big|_{H^s}\ \lesssim\ C_0\,M, \tag{4.4}$$

where we use that H^s is continuously embedded in $H^s_{\rm ul}$ and $H^s_{\rm ul}$ is a Banach algebra for all $s \geq s_0$ (e.g., Lannes 2013, Appendix B.4). The estimate on $V_{\rm rem}$, stated in the proposition, is given by (4.2)–(4.4).

Note that (4.4) yields in particular, for all $f \in H^s$, $s \geq s_0$, that

$$\begin{aligned} \big|\breve{m} f\big|_{H^s} \ \leq \ \big|\breve{m}\Lambda^s f\big|_{L^2} + \big|\big[\Lambda^s, \breve{m}\big] f\big|_{L^2} &\lesssim \big|\breve{m}\big|_{L^\infty}\big|f\big|_{H^s} \\ &\quad + \big|\partial_x \breve{m}\big|_{H^{\max\{s-1,s_0\}}}\big|f\big|_{H^{\max\{s-1,s_0\}}} \\ &\lesssim \ C_0\ M\ \big|f\big|_{H^s}, \end{aligned} \tag{4.5}$$

where we used the commutator estimate recalled in Lemma 5.3. Using the preceding estimates, it is now straightforward to check that $V^s_{\rm app} \equiv V_{\rm RL} + V^s_{\rm cor} + \varrho^2 V_{\rm rem} \equiv (\varrho\breve{\zeta}_1, \eta, v, \varrho^2\breve{m})^\top$ satisfies (3.3), up to a remainder term, R^s, with

$$\big\| R^s \big\|_{L^\infty([0,T];H^s)} \ \lesssim\ C_0\ M\ \varrho^2. \tag{4.6}$$

Here we used the fact that the occurrences of $\breve{m}$ in (3.3) are of the form $\partial_x\breve{m}$ or $\breve{m} \times f$, with $f \in H^s$, and both of these contributions are bounded in H^s, thanks to (4.4) and (4.5).

Construction of fast-mode approximate solution. The corrector $V^f_{\rm cor}$ has been defined as the unique solution to

$$\partial_t V^f_{\rm cor} \ + \ \frac{1}{\varrho} L_{(0)}\partial_x V^f_{\rm cor} \ = \ 0, \quad \text{where we recall } L_{(0)} \ \equiv \ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1+\delta^{-1} & 0 & 0 & 0 \end{pmatrix},$$

with initial data $V^f_{\rm cor}\,|_{t=0} \equiv (\zeta^0_1 - \varrho\breve{\zeta}_1\,|_{t=0}\,, 0, 0, m^0)^\top$.

Our aim is to prove that $V^f_{\rm cor}$ is an approximate solution to (3.3). Recall that the system reads

$$\partial_t V + \frac{1}{\varrho}\big(L_\varrho \ + \ \varrho B[V]\big)\,\partial_x V = 0, \quad \text{with}$$

$$L_\varrho \ \equiv \ \begin{pmatrix} 0 & 0 & \frac{\gamma-1}{\gamma(\delta+1)} & \frac{\gamma+\delta}{\gamma(\delta+1)} \\ 0 & 0 & \frac{\varrho}{1+\delta} & \frac{\varrho}{1+\delta} \\ 0 & \varrho(\gamma+\delta) & 0 & 0 \\ \gamma(1+\delta^{-1}) & \varrho\frac{\delta+\gamma}{\delta} & 0 & 0 \end{pmatrix}.$$

Thus $V^f_{\rm cor}$ satisfies

$$\partial_t V^f_{\rm cor} + \frac{1}{\varrho}\Big(L_\varrho \ + \ \varrho B[V^f_{\rm cor}]\Big)\,\partial_x V^f_{\rm cor} = R^f,$$

with

$$R^f \equiv \frac{1}{\varrho}(L_\varrho - L_{(0)})\partial_x V^f_{\rm cor} \;+\; B[V^f_{\rm cor}]\partial_x V^f_{\rm cor}.$$

It is obvious that for all $t \in \mathbb{R}$, $V^f_{\rm cor}$ satisfies

$$\big|V^f_{\rm cor}(t,\cdot)\big|_{X^{s+1}} \;\lesssim\; \big|V^f_{\rm cor}\,|_{t=0}\,\big|_{X^{s+1}} \;\le\; C_0\; M\; \varrho, \tag{4.7}$$

where we used (4.3) and the hypothesis on the initial data of the proposition.

In particular, Lemmas 3.7 and 5.1 yield

$$\big|B[V^f_{\rm cor}]\partial_x V^f_{\rm cor}\,\big|_{X^s} \;\lesssim\; \big|V^f_{\rm cor}\big|_{L^\infty(\mathbb{R})^4}\big|V^f_{\rm cor}\big|_{X^{s+1}} \;\le\; C_0\; M^2\; \varrho^2. \tag{4.8}$$

Now we use the fact that $(\mathrm{Id}-\Pi)V^f_{\rm cor} = V^f_{\rm cor}$, where we recall that Π represents the orthogonal projection onto $\ker(L_{(0)})$: $\mathrm{Id}-\Pi \equiv \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}$.

It is straightforward to check that

$$\big\|(L - L_{(0)})(\mathrm{Id}-\Pi)\big\| \;\lesssim\; \varrho^2,$$

so that

$$\Big|\frac{1}{\varrho}(L_\varrho - L_{(0)})\partial_x V^f_{\rm cor}\,\Big|_{X^s} \;=\; \Big|\frac{1}{\varrho}(L_\varrho - L_{(0)})(\mathrm{Id}-\Pi)\partial_x V^f_{\rm cor}\,\Big|_{X^s} \;\lesssim C_0\; M\; \varrho^2. \tag{4.9}$$

Estimates (4.8) and (4.9) immediately yield the desired result: $V^f_{\rm cor}$ satisfies (3.3), up to a remainder term, R^f, satisfying

$$\big\|R^f\big\|_{L^\infty([0,T];H^s)} \;\lesssim\; C_0\; M\; \varrho^2. \tag{4.10}$$

Completion of proof. One easily checks that $V_{\rm app} \equiv V^s_{\rm app} + V^f_{\rm cor}$ satisfies

$$\partial_t V_{\rm app} + \frac{1}{\varrho}\left(L_\varrho \;+\; \varrho B[V_{\rm app}]\right)\partial_x V_{\rm app} = R^f + R^s + R^c,$$

where

$$R^c \;\equiv\; (B[V_{\rm app}] - B[V^f_{\rm cor}])\partial_x V^f_{\rm cor} + (B[V_{\rm app}] - B[V^s_{\rm app}])\partial_x V^s_{\rm app}.$$

The contribution of $R^f + R^s$ is controlled as a result of the preceding calculations; see (4.6) and (4.10). Thus the only remaining term to control is R^c, which contains the coupling effects between $V^f_{\rm cor}$ and $V^s_{\rm app}$.

Note that, similarly to (3.20) in Lemma 3.7, one can check that estimates (4.2), (4.3), (4.4), (4.5), and (4.7) yield

$$\big|R^c\big|_{X^s} \;\leq\; C_0 \;\times\; \Big(\big\| V_{\mathrm{RL}} \otimes \partial_x V^f_{\mathrm{cor}} \big\|_{X^s} + \big\| V^f_{\mathrm{cor}} \otimes \partial_x V_{\mathrm{RL}} \big\|_{X^s} \;+\; M\,\varrho^2\Big), \qquad (4.11)$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$, and where $U \otimes V$ denotes the outer product of U and V.

To control the latter contribution, we make use of the fact that the initial data are assumed to be spatially localized. Thus V^f_{cor} is the superposition of two spatially localized waves, with center of mass $x \approx \pm c/\varrho t$. It follows that the contribution of the outer products will decay after some time, provided one can prove that V_{RL} remains spatially localized around $x = 0$ on the time interval $[0, T]$. This is where it is convenient, although certainly not necessary, to restrict ourselves to the time domain $t \in [0, T]$, with T bounded, instead of the more stringent $t \in [0, \tilde{T}/M]$. Indeed, as it roughly propagates with velocity ± 1, one cannot expect V_{RL} to remain spatially localized around $x = 0$ during the time interval $[0, T]$, with $T \gtrsim 1/M$, uniformly for M small.

We state and prove in what follows the persistence of the spatial decay that holds generically for a quasilinear, hyperbolic system and then complete the proof of Proposition 4.2.

Lemma 4.4 *(Persistence of spatial decay) Let $s \geq s_0 + 1,\ s_0 > 1/2$, and let $V_{\mathrm{RL}} \equiv (\eta, v)^\top$ be the solution to (1.2), with initial data $V_{\mathrm{RL}}\,|_{t=0} \equiv (\eta^0, v^0)^\top$ as previously. Assume moreover that there exists $\sigma > 0$ such that one has $\langle\cdot\rangle^\sigma \eta^0, \langle\cdot\rangle^\sigma v^0 \in H^s$ [where we use the notation $\langle x\rangle \equiv (1 + |x|^2)^{1/2}$]. There exists $M > 0$ such that if $\big|(\eta^0, v^0)^\top\big|_{H^s \times H^s} \leq M$, then one has*

$$\begin{aligned} \forall t \in [0, T], \qquad & \big|\langle\cdot\rangle^\sigma \eta\big|_{H^s} + \big|\langle\cdot\rangle^\sigma v\big|_{H^s} \\ & \leq\; C\big(M, h_0^{-1}, \big|\langle\cdot\rangle^\sigma \eta^0\big|_{H^s} + \big|\langle\cdot\rangle^\sigma v^0\big|_{H^s}, \delta_{\min}^{-1}, \delta_{\max}\big). \end{aligned}$$

Proof of Lemma Consider $W(t, x) = \langle x\rangle^\sigma V_{\mathrm{RL}}(t, x)$ (here and in what follows, multiplying a vector-valued function by $\langle x\rangle^\sigma$ means that all components are multiplied). One has

$$S[V_{\mathrm{RL}}]\partial_t\big(\langle\cdot\rangle^{-\sigma} W\big) + \Sigma[V_{\mathrm{RL}}]\partial_x\big(\langle\cdot\rangle^{-\sigma} W\big) = 0,$$

where $S[\cdot]$, $\Sigma[\cdot]$ are smooth mappings onto the space of 2×2 symmetric matrices (S and Σ are explicit; see Guyenne et al. 2010 for more details).

It follows, since the multiplication by $\langle\cdot\rangle^\sigma$ obviously commutes with $S[\cdot]$, $\Sigma[\cdot]$, ∂_t, that

$$S[V_{\mathrm{RL}}]\partial_t W + \Sigma[V_{\mathrm{RL}}]\partial_x W + \langle x\rangle^\sigma \partial_x\big(\langle x\rangle^{-\sigma}\big)\Sigma[V_{\mathrm{RL}}]W = 0.$$

$S[V_{\mathrm{RL}}]$ is positive definite, so that there exists $0 < c_0 < \infty$ such that

$$\frac{1}{c_0}\big|W\big|^2_{H^s(\mathbb{R})^2} \;\leq\; E^s(W) \equiv \big(S[V_{\mathrm{RL}}]\Lambda^s W, \Lambda^s W\big) \;\leq\; c_0\big|W\big|^2_{H^s(\mathbb{R})^2}.$$

Using the usual technique for a priori H^s estimates (e.g., Lemma 5.6) we obtain

$$\frac{d}{dt}E^s(W) \le C\big(\big|V_{\mathrm{RL}}\big|_{X^s}, \big|\partial_t V_{\mathrm{RL}}\big|_{X^{s-1}}\big)\ E^s(W) \\ + C\big(\big|\langle x\rangle^\sigma \partial_x\big(\langle x\rangle^{-\sigma}\big)\big|_{H^s}, \big|V_{\mathrm{RL}}\big|_{X^s}\big)\ E^s(W)^{1/2}.$$

Now, using the control of $V_{\mathrm{RL}} \in X^s$ in (4.2), and since we have

$$\big|\langle x\rangle^\sigma \partial_x\big(\langle x\rangle^{-\sigma}\big)\big|_{H^s} = \big|\sigma x\langle x\rangle^{-2}\big|_{H^s} \lesssim \sigma,$$

it follows from the Gronwall–Bihari inequality that

$$E^s(W) \ \le\ E^s(W\,|_{t=0})\exp(C_0 t) \ +\ \int_0^t C_1\exp(C_0(t-t'))\,dt',$$

with $C_0, C_1 = C\big(M, h_0^{-1}, \big|\langle\cdot\rangle^\sigma\eta^0\big|_{H^s} + \big|\langle\cdot\rangle^\sigma v^0\big|_{H^s}, \delta_{\min}^{-1}, \delta_{\max}\big)$, and the lemma is proved. □

Let us now complete the proof of Proposition 4.2. We use the following calculation to estimate R^c in (4.11). Set $s > 1/2$, $\sigma > 0$, and $c \neq 0$. Let u, v satisfy $\langle\cdot\rangle^\sigma v(t,\cdot) \in H^s$, and $\langle\cdot\rangle^\sigma u(\cdot) \in H^s$. Then we have

$$\big|v(\cdot)u_\pm(\cdot - c/\varrho t)\big|_{H^s} \\ \lesssim \big|(1+|\cdot|^2)^\sigma v\big|_{H^s}\big|(1+|\cdot|^2)^\sigma u\big|_{H^s}\big|(1+|\cdot|^2)^{-\sigma}(1+|\cdot - c/\varrho t|^2)^{-\sigma}\big|_{H^s},$$

and we can check (e.g., Lannes (2003)) that for all $\sigma > 1/2$ and $T > 0$ we have

$$\int_0^T \big|(1+|\cdot|^2)^{-\sigma}(1+|\cdot - c/\varrho t'|^2)^{-\sigma}\big|_{H^s}\,dt' \ \le\ C\big(\frac{1}{2\sigma-1}, \frac{1}{c}\big)\,\varrho,$$

thus uniformly bounded with respect to $1/\varrho$ and T.

It is now straightforward, applying Lemma 4.4, the definition of V^f_{cor}, (4.7), and the preceding calculations to (4.11), that the following estimate holds:

$$\int_0^T \big|R^c(t',\cdot)\big|_{X^s}\,dt' \ \le\ C_0\,M\,\varrho^2, \tag{4.12}$$

with $C_0 = C(M, h_0^{-1}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$.

Estimates (4.6), (4.10), and (4.12) complete the proof of Proposition 4.2. □

Let us conclude this section with the following result, which corresponds to Theorem 1.2 when Proposition 4.2 is used instead of Proposition 3.4.

Theorem 4.5 *Let $s \geq s_0 + 1, s_0 > 1/2$. Let $\zeta_1^0, \zeta_2^0, u_1^0, u_2^0 \in H^{s+1}(\mathbb{R})$ be such that* (1.6) *holds with $h_0 > 0$, and there exists $0 < M < \infty$ and $\sigma > 1/2$ such that*

$$\big|(1+|\cdot|^2)^\sigma \zeta_2^0\big|_{H^{s+1}} + \big|(1+|\cdot|^2)^\sigma (u_2^0 - \gamma u_1^0)\big|_{H^{s+1}} \leq M \tag{4.13}$$

and

$$\big|(1+|\cdot|^2)^\sigma \zeta_1^0\big|_{H^{s+1}} + \big|(1+|\cdot|^2)^\sigma \big(\gamma h_1^0 u_1^0 + h_2^0 u_2^0\big)\big|_{H^{s+1}} \leq M\varrho. \tag{4.14}$$

Then there exists T^{-1}, C, depending nondecreasingly on $M, h_0^{-1}, \frac{1}{s_0-1/2}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}$, $\delta_{\max}, \gamma_{\min}^{-1}$, such that one can uniquely define $U \in C([0,T]; X^{s+1}) \cap C^1([0,T]; X^s)$, the solution to (1.1) *with initial data $U|_{t=0} = (\zeta_1^0, \zeta_2^0, u_1^0, u_2^0)^\top$, and $V_{\mathrm{RL}}, V_{\mathrm{cor}}^s, V_{\mathrm{cor}}^f$ as in Proposition 4.2. Denote by U_{app} the approximate solution corresponding to $V_{\mathrm{RL}} + V_{\mathrm{cor}}^s + V_{\mathrm{cor}}^f$, after the change of variables in* (3.2). *Then one has*

$$\big\|U - U_{\mathrm{app}}\big\|_{L^\infty([0,T];X^s_{\mathrm{ul}})} \leq C\, M\, \varrho^2.$$

Sketch of the proof The existence and uniqueness of U were stated in Theorem 1.2. The existence and uniqueness of $V_{\mathrm{RL}}, V_{\mathrm{cor}}^s, V_{\mathrm{cor}}^f$ are guaranteed by Proposition 4.2. Now we can follow the same procedure as described in Sect. 3 (and especially Sect. 3.3) using the result of Proposition 4.2 instead of the corresponding Proposition 3.4. Note, however, that the remainder term constructed in Proposition 4.2, V_{rem}, may not have a finite H^s norm; thus we need to work with uniformly local Sobolev spaces, defined in Remark 4.3.

However, as initially remarked by Kato (1975), the energy method for hyperbolic quasilinear systems in Sobolev spaces extends naturally to uniformly local Sobolev spaces, without significant changes in the proof (in particular, similar product and commutator estimates hold; see Lannes (2013), Appendix B; thus we do not go further into details.

We simply remark that V_{app} was constructed so that $W \equiv V - V_{\mathrm{app}}$ satisfies

$$\big|W|_{t=0}\big|_{X^s_{\mathrm{ul}}} \lesssim C_0 M \varrho^2,$$

where we denote by $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top$ the solution to (3.3) corresponding to U, in terms of the variables defined by (3.1). Consequently, the energy estimate (3.26) in Lemma 3.9 implies

$$\forall t \in [0,T], \quad \big|W\big|_{X^s_{\mathrm{ul}}} \lesssim C_0 M \varrho^2 + \int_0^t \big|R(t',\cdot)\big|_{X^s_{\mathrm{ul}}}\, dt',$$

and Proposition 4.2 immediately yields the desired estimate. □

4.2 The Case of Ill-Prepared Initial Data

In this section, we are concerned with the case of ill-prepared initial data, that is, initial data that fail to meet the smallness assumption in (1.5) or, in other words, that *admit a nonsmall fast mode*. Once again, we construct an approximate solution as the superposition of a slow-mode approximate solution, obtained from the corresponding solution to the rigid-lid system (1.2), and a fast-mode approximate solution, which we will present subsequently. There are two main differences from the previous results due to the fact that the slow-mode approximate solution is no longer of size $\mathcal{O}(\varrho)$:

1. Nonlinear effects have a nontrivial effect on the behavior of the fast-mode approximate solution and cannot be neglected.
2. The strategy developed in Sect. 3 is no longer valid because the hypothesis of Lemma 3.9 is no longer satisfied.

As a consequence of the latter point, we restrict our statement to a consistency result, namely Proposition 4.6 (below); we cannot deduce an estimate on the difference between the exact and approximate solutions, as in Theorems 1.2 and 4.5, or even prove that (1.1) is well-posed on a time interval independent of small ϱ. However, numerical simulations presented in the next subsection are in full agreement with the intuitive conjecture that

$$\big\| V - V_{\mathrm{RL}} - V_{\mathrm{cor}}^{f} \big\|_{L^\infty([0,T];X^s)} \;=\; \mathcal{O}(\varrho),$$

with the notations introduced below.

Proposition 4.6 *Let $s \geq s_0$, $s_0 > 1/2$, and $\zeta_1^0, \zeta_2^0, u_s^0, m^0 \in H^{s+1}(\mathbb{R})$, satisfying (1.6) [following the change of variable in (3.2)] with given $h_0 > 0$. Assume additionally that there exists $0 < M < \infty$ and $\sigma > 1/2$ such that*

$$\begin{aligned}&\big|(1+|\cdot|^2)^\sigma \zeta_1^0\big|_{H^{s+2}} + \big|(1+|\cdot|^2)^\sigma m^0\big|_{H^{s+2}} + \big|(1+|\cdot|^2)^\sigma \zeta_2^0\big|_{H^{s+2}}\\ &\quad + \big|(1+|\cdot|^2)^\sigma u_s^0\big|_{H^{s+2}} \;\leq\; M.\end{aligned}$$

Then there exists $0 < T^{-1}, C_0 \leq C(M, h_0^{-1}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$ such that

1. *$V_{\mathrm{RL}} \equiv (0, \eta, v, 0)^\top$ is well defined by Definition 3.3 and satisfies*

$$\forall t \in [0,T], \qquad \big|V_{\mathrm{RL}}\big|_{X^{s+2}} + \big|\partial_t V_{\mathrm{RL}}\big|_{X^{s+1}} \;\leq\; C_0\, M;$$

2. *V_{cor}^{f} is well defined with*

$$V_{\mathrm{cor}}^{f}(t,x) \equiv \begin{pmatrix} u_+(t,x) + u_-(t,x) \\ 0 \\ 0 \\ c(u_+(t,x) - u_-(t,x)) \end{pmatrix},$$

where $c \equiv \sqrt{1+\delta^{-1}}$, *and* $u_\pm$ *is the unique solution to*

$$\partial_t u_\pm \ \pm\ \frac{c}{\varrho}\partial_x u_\pm \ \pm\ \frac{3}{2c}u_\pm \partial_x u_\pm \ =\ 0,$$

with $u_\pm\,|_{t=0} \ =\ \frac{1}{2}\big(\zeta_1^0 \pm c^{-1}m^0\big)$;

3. *There exists* $V_{\rm rem}$ *with*

$$\forall t \in [0,T], \qquad \big|V_{\rm rem}\big|_{X^{s+1}} + \varrho\big|\partial_t V_{\rm rem}\big|_{X^s} \ \le\ C_0\, M$$

such that $V_{\rm app} \equiv V_{\rm RL} + V_{\rm cor}^f + \varrho V_{\rm rem}$ *satisfies* (3.3)*, up to a remainder term, R, with*

$$\int_0^T \big|R(t,\cdot)\big|_{X^s}\, dt \ \le\ C_0\, M\, \varrho.$$

Remark 4.7 The fast-mode contribution $V_{\rm cor}^f$ is different from that defined in Proposition 4.2. Moreover, it is not a corrector term *per se* since it has the same order of magnitude as $V_{\rm RL}$. We decided to use the same notation in order to acknowledge the following fact: one can replace $V_{\rm cor}^f$ in Proposition 4.2 by the one defined previously without modifying the rest of the statement; nonlinear effects on the fast-mode component are negligible in the case of well-prepared initial data.

Proof of Proposition 4.6 We follow the same three steps as in the proof of Proposition 4.2. We first construct an approximate solution corresponding to the slow mode and fast mode. Finally, we prove that the coupling effects between the two modes are weak due to the appropriate spatial localization of the initial data, and therefore the superposition of the two modes yields an approximate solution.

Construction of slow-mode approximate solution. Proposition 3.4 directly gives the desired result: using the notation $V_{\rm rem}^s \equiv (\breve\zeta_1, 0, 0, \varrho\breve m)$, with $\breve\zeta_1, \breve m$ as defined in (3.14), (3.16), we have

$$\forall t \in [0,T], \qquad \big|V_{\rm RL}\big|_{X^{s+2}} + \big|\partial_t V_{\rm RL}\big|_{X^{s+1}} \ \lesssim\ C_0\, M, \tag{4.15}$$

$$\forall t \in [0,T], \qquad \big|V_{\rm rem}^s\big|_{X^{s+2}} + \big|\partial_t V_{\rm rem}^s\big|_{X^{s+1}} \ \lesssim\ C_0\, M, \tag{4.16}$$

and $V_{\rm app}^s \equiv V_{\rm RL} + \varrho V_{\rm rem}^s$ satisfies (3.3) up to a remainder term, R^s, with

$$\big\|R^s\big\|_{L^\infty([0,T];X^{s+1})} \ \lesssim\ C_0\, M(M\,\varrho + \varrho^2) \lesssim\ C_0\, M\, \varrho\,, \tag{4.17}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$. As previously, the first steps of the proof are valid with $T = \tilde T/M$, but the last step—because it uses the localization in space of the two modes—requires that T be uniformly bounded.

Construction of fast-mode approximate solution. We recall that (3.3) reads

$$\partial_t V + \frac{1}{\varrho}\left(L_\varrho \; + \; \varrho B[V]\right)\partial_x V = 0,$$

with $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top$. We use the notation $L_\varrho \;\equiv\; L_{(0)} + \varrho L_{(1)} + \mathcal{O}(\varrho^2)$, with

$$L_{(0)} \;\equiv\; \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1+\delta^{-1} & 0 & 0 & 0 \end{pmatrix}, \quad L_{(1)} \;\equiv\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{1+\delta} & \frac{1}{1+\delta} \\ 0 & \gamma+\delta & 0 & 0 \\ 0 & \frac{\delta+1}{\delta} & 0 & 0 \end{pmatrix}.$$

We can also check that $B[(\zeta_1, 0, 0, m)^\top] \equiv B_{(1)}[(\mathrm{Id} - \Pi)V] + \mathcal{O}(\varrho)$, with

$$B_{(1)}[(\mathrm{Id} - \Pi)V] \;\equiv\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\delta}{\delta+1}m & 0 & 0 \\ 0 & 0 & \frac{\delta}{\delta+1}m & 0 \\ \zeta_1 & 0 & 0 & 2\frac{\delta}{\delta+1}m \end{pmatrix}.$$

In what follows, we seek an approximate solution to

$$\partial_t V + \left(\frac{1}{\varrho}L_{(0)} \; + \; L_{(1)} \; + \; B_{(1)}[(\mathrm{Id} - \Pi)V]\right)\partial_x V = 0, \tag{4.18}$$

with initial data satisfying $(\mathrm{Id} - \Pi)V\,|_{t=0} \;=\; V\,|_{t=0}$.

Our strategy is based on a WKB-type expansion, that is, we seek an approximate solution to (4.18) under the form

$$V^f_{\mathrm{app}}(t,x) \;=\; V^f_{\mathrm{cor}}(t, t/\varrho, x) \; + \; \varrho V^f_{\mathrm{rem}}(t, t/\varrho, x),$$

where (with a straightforward abuse of notation) $V^f_{\mathrm{app}}(t,\tau,x)$ is an approximate solution to

$$\frac{1}{\varrho}\partial_\tau V^f_{\mathrm{app}} + \partial_t V^f_{\mathrm{app}} + \left(\frac{1}{\varrho}L_{(0)} \; + \; L_{(1)} \; + \; B_{(1)}[(\mathrm{Id} - \Pi)V^f_{\mathrm{app}}]\right)\partial_x V^f_{\mathrm{app}} = 0. \tag{4.19}$$

Based on the fact that at first order (in terms of ϱ) the system (4.19) is a simple linear equation, $\partial_\tau V \; + \; L_{(0)}\partial_x V \;=\; 0$, and from the assumption on the initial data we set V^f_{cor} as the superposition of decoupled waves, supported on the eigenvectors of $L_{(0)}$ corresponding to nonzero eigenvalues.

The analysis of higher-order terms yields

- the behavior of V^f_{cor} with respect to the large-time-scale variable, t, which takes into account the nonlinear effects on the propagation of each decoupled wave;
- a remainder term, $V^f_{\mathrm{rem}}(t,\tau,x)$, that mimics the coupling effects between the two counterpropagating waves of V^f_{cor}, as well as the "slow-mode component," ΠV^f_{cor}.

The key ingredient in the proof is to show that one can set $V^f_{\rm cor}$ such that $V^f_{\rm rem}$ remains small for large time scales. This strategy has been applied notably to the rigorous justification of the Korteweg—de Vries equation as a model for the propagation of surface waves in the long wave regime (Schneider and Wayne 2000; Bona et al. 2005) and subsequently to similar problems in the bifluidic setting (Duchêne 2011, 2014). The strategy is described comprehensively in, for example, Lannes (2013), Chap. 7; thus we do not provide details of the calculations but simply state the outcome.

It is convenient to introduce here the following eigenvectors of $L_{(0)}$:[4]

$$\mathbf{e}_+ = \begin{pmatrix}1\\0\\0\\c\end{pmatrix}, \quad \mathbf{e}_- = \begin{pmatrix}1\\0\\0\\-c\end{pmatrix}, \quad \mathbf{e}_0 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}.$$

We set

$$V^f_{\rm cor}(\cdot,\tau,x) \equiv u_+(\cdot, x-c\tau)\mathbf{e}_+ \;+\; u_-(\cdot,x+c\tau)\mathbf{e}_-,$$

where $u_\pm(t,y)$ is uniquely defined by

$$\partial_t u_\pm \;\pm\; \frac{3}{2c}u_\pm\partial_y u_\pm \;=\; 0,$$

with $u_\pm|_{t=0} \;=\; \frac{1}{2}\big(\zeta_1^0 \pm c^{-1}m^0\big)$. We check immediately that $V^f_{\rm cor} \;:\; (t,x) \mapsto V^f_{\rm cor}(t,t/\varrho,x)$ is as in the proposition, explaining our (slightly misused) notation.

In the same way, we write

$$V^f_{\rm rem}(\cdot,\tau,x) \equiv r_+(t,\tau,x)\mathbf{e}_+ + r_-(t,\tau,x)\mathbf{e}_- + r_0(t,\tau,x)\mathbf{e}_0,$$

with functions r_+, r_-, r_0 determined by

$$\begin{aligned}
&\partial_\tau r_+(\cdot,\tau,x) + c\partial_x r_+(\cdot,\tau,x) + \frac{3}{4c}\partial_x\big(u_-(\cdot,x-c\tau)^2\big)\\
&\quad -\;\frac{1}{2c}\partial_x\big(u_-(\cdot,x-c\tau)u_+(\cdot,x+c\tau)\big) \;=\; 0,\\
&\partial_\tau r_-(\cdot,\tau,x) - c\partial_x r_+(\cdot,\tau,x) - \frac{3}{4c}\partial_x\big(u_+(\cdot,x-c\tau)^2\big)\\
&\quad +\;\frac{1}{2c}\partial_x\big(u_-(\cdot,x-c\tau)u_+(\cdot,x+c\tau)\big) \;=\; 0,\\
&\partial_\tau r_0(\cdot,\tau,x) \;+\; \frac{1}{\delta c}\partial_x\big(u_+(\cdot,x+c\tau) - u_-(\cdot,x-c\tau)\big) \;=\; 0,
\end{aligned}$$

and $V^f_{\rm rem}(\cdot,0,\cdot) \equiv 0$.

[4] Of course a fourth vector—second linearly independent element of $\ker(L_{(0)})$—could be defined, but this is not necessary in our analysis.

We can check that $V^f_{\rm app}(t,\tau,x) = V^f_{\rm cor}(t,\tau,x) + \varrho V^f_{\rm rem}(t,\tau,x)$, as defined previously, satisfies

$$\frac{1}{\varrho}\partial_\tau V^f_{\rm app} + \partial_t V^f_{\rm app} + \Big(\frac{1}{\varrho}L_{(0)} + L_{(1)} + B_{(1)}[V^f_{\rm app}]\Big)\partial_x V^f_{\rm app} = R^f,$$

with $R^f \equiv \varrho\partial_t V^f_{\rm rem} + \varrho L_{(1)}\partial_x V^f_{\rm rem} + B_{(1)}[V^f_{\rm app}]\partial_x V^f_{\rm app} - B_{(1)}[V^f_{\rm cor}]\partial_x V^f_{\rm cor}$.

It follows [using (3.20) in Lemma 3.7] that

$$\big|R^f\big|_{X^s} \le \varrho\, C\big(\big|\partial_t V^f_{\rm rem}\big|_{X^s}, \big|V^f_{\rm rem}\big|_{X^{s+1}}, \big|V^f_{\rm cor}\big|_{X^{s+1}}\big). \tag{4.20}$$

To estimate the preceding quantity, one needs to control $V^f_{\rm rem}$ using the following two lemmata.

Lemma 4.8 *Let $s \ge 0$ and $f^0 \in H^s(\mathbb{R})$. Then there exists a unique global strong solution, $u(\tau,x) \in C^0(\mathbb{R}; H^s) \cap C^1(\mathbb{R}; H^{s-1})$, of*

$$\begin{cases}(\partial_\tau + c_1\partial_x)u = \partial_x f\\ u\,|_{t=0} = 0\end{cases} \quad \textit{with} \quad \begin{cases}(\partial_\tau + c_2\partial_x)f = 0,\\ f_i\,|_{t=0} = f^0,\end{cases}$$

where $c_1 \neq c_2$. Moreover, we have the following estimates for all $\tau \in \mathbb{R}$:

$$\big|u(\tau,\cdot)\big|_{H^s(\mathbb{R})} \le \frac{2}{|c_1 - c_2|}\big|f^0\big|_{H^s(\mathbb{R})}.$$

Lemma 4.9 *Let $s \ge s_0 > 1/2$ and $v_1^0, v_2^0 \in H^s(\mathbb{R})$. Then there exists a unique global strong solution, $u \in C^0(\mathbb{R}; H^s)$, of*

$$\begin{cases}(\partial_\tau + c\partial_x)u = g(v_1, v_2)\\ u\,|_{t=0} = 0\end{cases} \quad \textit{with} \quad \forall i \in \{1,2\} \quad \begin{cases}(\partial_\tau + c_i\partial_x)v_i = 0,\\ v_i\,|_{t=0} = v_i^0,\end{cases}$$

where $c_1 \neq c_2$ and g is a bilinear mapping defined on $\mathbb{R}^2$ and with values in $\mathbb{R}$. Assume, moreover, that there exists $\sigma > 1/2$ such that $v_1^0(1+|\cdot|^2)^\sigma$ and $v_2^0(1+|\cdot|^2)^\sigma \in H^s(\mathbb{R})$; then one has the (uniform in time) estimate

$$\big\|u\big\|_{L^\infty(\mathbb{R};H^s(\mathbb{R}))} \le C\Big(\frac{1}{c_1 - c_2}, \frac{1}{\sigma - 1/2}\Big)\, \big|v_1^0(1+|\cdot|^2)^\sigma\big|_{H^s(\mathbb{R})}\big|v_2^0(1+|\cdot|^2)^\sigma\big|_{H^s(\mathbb{R})}.$$

Lemma 4.8 is straightforward, and Lemma 4.9 follows from Proposition 3.5 in Lannes (2003).

Lemmata 4.8 and 4.9 applied to $V^f_{\rm rem}$ immediately yield

$$\begin{aligned}\big|V^f_{\rm rem}(t,\tau,\cdot)\big|_{X^{s+1}} &\le C\big|u_\pm(t,\cdot)\big|_{H^{s+1}(\mathbb{R})}\\ &\quad + C\big|u_+(t,\cdot)(1+|\cdot|^2)^\sigma\big|_{H^{s+2}(\mathbb{R})}\big|u_-(t,\cdot)(1+|\cdot|^2)^\sigma\big|_{H^{s+2}(\mathbb{R})}.\end{aligned}$$

We can apply the same arguments to $\partial_t V^f_{\rm rem}$ [differentiating the equations satisfied by $r_\pm$ and r_0 with respect to the parameter t and using $\partial_t u_\pm(t,y) = \mp\frac{3}{2c}u_\pm(t,y)\partial_y u_\pm(t,y)$] and obtain

$$\begin{aligned}\big|\partial_t V^f_{\rm rem}(t,\tau,\cdot)\big|_{X^s} \;&\leq\; C\big|u_\pm(t,\cdot)\big|_{H^{s+1}(\mathbb{R})}\\ &+ C\big|u_+(t,\cdot)(1+|\cdot|^2)^\sigma\big|_{H^{s+2}(\mathbb{R})}\big|u_-(t,\cdot)(1+|\cdot|^2)^\sigma\big|_{H^{s+2}(\mathbb{R})}.\end{aligned}$$

It is not difficult to show that the inviscid Burgers equation propagates locally in time the localization in space of its solutions (Lemma 4.4), so that we have

$$\forall t\in[0,T],\quad \big|u_\pm(t,\cdot)(1+|\cdot|^2)^\sigma\big|_{H^{s+2}(\mathbb{R})} \;\lesssim\; \big|u_\pm(0,\cdot)(1+|\cdot|^2)^\sigma\big|_{H^{s+2}(\mathbb{R})} \leq M, \tag{4.21}$$

and thus we have proved

$$\forall(t,\tau)\in[0,T]\times\mathbb{R},\quad \big|V^f_{\rm rem}(t,\tau,\cdot)\big|_{X^{s+1}} + \big|\partial_t V^f_{\rm rem}(t,\tau,\cdot)\big|_{X^s} \;\leq\; C_0\,M,$$

with $C_0 = C(M,\frac{1}{2\sigma-1},\delta_{\min}^{-1},\delta_{\max},\gamma_{\min}^{-1})$.

Finally, we recall that $V^f_{\rm cor} \equiv u_+(t,x-ct/\varrho)\mathbf{e}_+ + u_-(t,x+ct/\varrho)\mathbf{e}_-$ and $V^f_{\rm rem} \equiv V^f_{\rm rem}(t,t/\varrho,x)$ and deduce

$$\forall t\in[0,T],\qquad \big|V^f_{\rm cor}\big|_{X^{s+2}} + \varrho\big|\partial_t V^f_{\rm cor}\big|_{X^{s+1}} \;\leq\; C_0\,M, \tag{4.22}$$

$$\forall t\in[0,T],\qquad \big|V^f_{\rm rem}\big|_{X^{s+1}} + \varrho\big|\partial_t V^f_{\rm rem}\big|_{X^{s}} \;\leq\; C_0\,M, \tag{4.23}$$

with $C_0 = C(M,\frac{1}{2\sigma-1},\delta_{\min}^{-1},\delta_{\max},\gamma_{\min}^{-1})$.

Therefore, (4.20) simply becomes

$$\big\|R^f\big\|_{L^\infty([0,T];X^s)} \;\leq\; C_0\,M\,\varrho, \tag{4.24}$$

with $C_0 = C(M,\frac{1}{2\sigma-1},\delta_{\min}^{-1},\delta_{\max},\gamma_{\min}^{-1})$.

Completion of proof. We can easily check that $V_{\rm app} \equiv V^s_{\rm app} + V^f_{\rm app} \equiv V_{\rm RL} + V^f_{\rm cor} + \varrho V^s_{\rm rem} + \varrho V^f_{\rm rem}$ satisfies

$$\partial_t V_{\rm app} + \frac{1}{\varrho}\big(L_\varrho \;+\; \varrho B[V_{\rm app}]\big)\,\partial_x V_{\rm app} = R^s + R^s + R^c,$$

where R^s and R^f are as defined and estimated previously, and with

$$R^c \;\equiv\; (B[V_{\rm app}] - B[V^f_{\rm app}])\partial_x V^f_{\rm app} + (B[V_{\rm app}] - B[V^s_{\rm app}])\partial_x V^s_{\rm app}.$$

The contribution of $R^f + R^s$ is controlled as a result of the preceding calculations; see (4.17) and (4.24). Thus the only component to control comes from the coupling effects between $V^s_{\rm app}$ and $V^f_{\rm app}$, presented in R^c. Recalling the construction of $V^s_{\rm app} \equiv$

$V_{\rm RL} + \varrho V^s_{\rm rem}$ and $V^f_{\rm app} \equiv V^f_{\rm cor} + \varrho V^f_{\rm rem}$ and using estimates (4.15), (4.16), (4.22), and (4.23), we can check that

$$\big|R^c\big|_{X^s} \ \leq\ C_0 \times \Big(\big\| V_{\rm RL} \otimes \partial_x V^f_{\rm cor}\big\|_{X^s} + \big\| V^f_{\rm cor} \otimes \partial_x V_{\rm RL}\big\|_{X^s} \ +\ M\varrho\Big),$$

with $C_0 = C(M, h_0^{-1}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$, and again $U \otimes V$ is the outer product of U and V.

We estimate the preceding expression as in the proof of Proposition 4.2 using spatial localization. For any function v satisfying $(1 + |\cdot|^2)^\sigma v(t,\cdot) \in H^s$ we have

$$\begin{aligned}&\big|v(t,\cdot)u_\pm(t,\cdot \mp c/\varrho t)\big|_{H^s}\\ &\quad\lesssim \big|(1+|\cdot|^2)^\sigma v(t,\cdot)\big|_{H^s}\big|(1+|\cdot|^2)^\sigma u_\pm(t,\cdot)\big|_{H^s}\big|(1+|\cdot|^2)^{-\sigma}(1+|\cdot \mp c/\varrho t|^2)^{-\sigma}\big|_{H^s},\end{aligned}$$

and we recall that for all $\sigma > 1/2$ and $t > 0$, we have

$$\int_0^t \big|(1+|\cdot|^2)^{-\sigma}(1+|\cdot \mp c/\varrho t'|^2)^{-\sigma}\big|_{H^s}\, dt' \ \leq\ C\Big(\frac{1}{2\sigma-1}, \frac{1}{c}\Big)\, M\,\varrho,$$

thus uniformly bounded with respect to $1/\varrho$ and T.

Hence it follows from Lemma 4.4 and (4.21) that one can restrict $T > 0$ such that

$$\int_0^T \big|R^c(t,\cdot)\big|_{X^s}\, dt \ \leq\ \ C_0\, M\, \varrho,$$

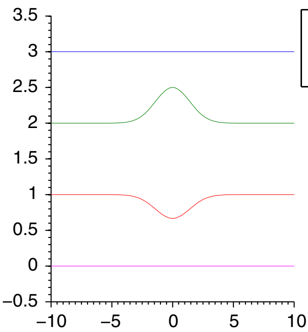
with $C_0 = C(M, h_0^{-1}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$. Proposition 4.6 is proved. □

Remark 4.10 As mentioned previously, we are unable to deduce from Proposition 4.6 a rigorous estimate on the difference between the exact solution and the constructed approximate solution as in Theorem 1.2 or 4.5. Indeed, the strategy developed in Sect. 3.3 fails because the solution does not satisfy the assumption of Lemma 3.9 or, more precisely, the estimate on the time derivative, $\partial_t V$. A closer look at the proof shows that the only problematic term to estimate is $\big|\big[\partial_t, T[V]\big]\Lambda^s W\big|_{L^2}$ or, even more precisely, $\big|\big[\partial_t, T[V]\big]\Pi\Lambda^s W\big|_{L^2}$ because the supplementary is estimated through (3.24) in Lemma 3.8. We expect that the following strategy would imply the desired result: decompose

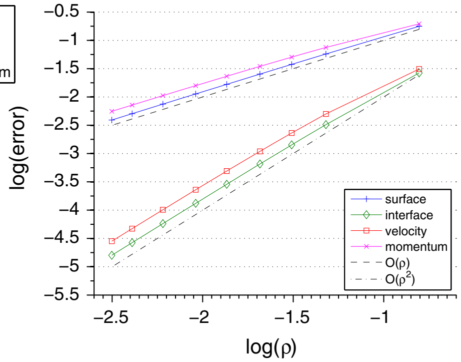
$$\big|\big[\partial_t, T[V]\big]\Pi\Lambda^s W\big|_{L^2} \lesssim \big\|(\Pi\Lambda^s W)\otimes \Pi\partial_t V\big\|_{L^2} + \big\|(\Pi\Lambda^s W)\otimes ({\rm Id}-\Pi)\partial_t V\big\|_{L^2}.$$

The first term is uniformly bounded because $\Pi\partial_t V$ roughly corresponds to the slow mode of the flow; the second term can be estimated using the different spatial localization of ΠW and $({\rm Id}-\Pi)V$.

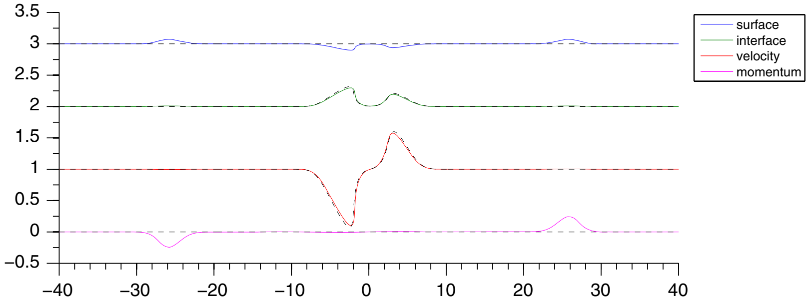
Following this strategy would require a few technical results and lengthy calculations, and hence we do not pursue it. Let us simply remark that the numerical simulations presented in the following section show perfect agreement with the desired



(a) Initial data



(b) Error with respect to ϱ (log-log scale)



(c) Flow at finite time $T = 4$

Fig. 2 Solution of free-surface system compared with rigid-lid approximate solution

result, namely

$$\left\| V - V_{\mathrm{RL}} - V_{\mathrm{cor}}^{f} \right\|_{L^\infty([0,T];X^s)} = \mathcal{O}(\varrho).$$

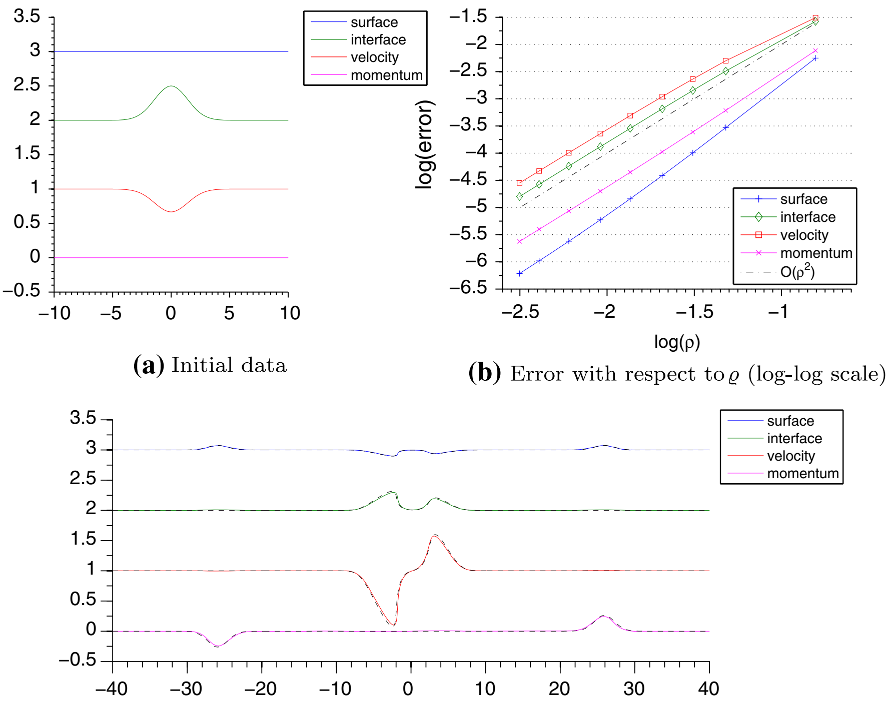
4.3 Discussion and Numerical Simulations

In this section, we illustrate and discuss the results presented in Theorem 1.2 and Proposition 3.4 (validity of rigid-lid approximation), Proposition 4.2 and Theorem 4.5 (improved approximate solution), and Proposition 4.6 (case of ill-prepared initial data).

In each case, we construct the appropriate approximate solution $(V_{\mathrm{RL}}, V_{\mathrm{cor}}^{f}, V_{\mathrm{cor}}^{s})$ and compare with the exact solution of the free-surface system (3.3) [which is equivalent to (1.1) with the corresponding variables] for different values of ϱ (and $\alpha = \varrho$) while the other parameters are fixed.

More precisely, we set

$$\delta = 1/2\ ;\ \ \epsilon = 1/2;$$
$$\gamma \in \{0.75, 0.9, 0.93, 0.95, 0.965, 0.0975, 0.09825, 0.09875, 0.099\}.$$



(a) Initial data **(b)** Error with respect to ϱ (log-log scale)

(c) Flow at finite time $T = 4$

Fig. 3 Solution of free-surface system compared with improved approximate solution

The initial data are set as follows:

$$\zeta_2\mid_{t=0} = \exp\big(-(x/2)^2\big) \quad ; \quad u_s\mid_{t=0} = \frac{-1}{3}\exp\big(-(x/2)^2\big),$$

and

$$\zeta_1\mid_{t=0} = 0 \quad ; \quad u_s\mid_{t=0} = \begin{cases} 0 & \text{in the well-prepared case;} \\ 2\exp\big(-(x/2)^2\big) & \text{in the ill-prepared case.} \end{cases}$$

We compute for times $t \in [0, T]$, with $T = 4$.

Each figure contains three panels. The upper-left panel represents the initial data. For the sake of readability, we plot respectively $1+\delta^{-1}+\epsilon\zeta_1\mid_{t=0}$, $\delta^{-1}+\epsilon\zeta_2\mid_{t=0}$, $1+u_s\mid_{t=0}$, and $m\mid_{t=0}$. The lower panel represents the solution of the free-surface system (3.3) and the corresponding approximate solution of interest (the latter with dotted lines), at final time $T = 4$, for $\gamma = 0.9$, thus $\varrho \approx 0.2673$. Finally, in the upper-right panel, we plot the normalized discrete l^2-norm of the difference between the aforementioned data in a loglog scale, for several values of ϱ (the markers reveal the positions that have been computed), at final time $T = 4$.

The numerical scheme we use takes advantage of spectral methods for the space discretization (see Trefethen 2000) and thus yields an exponential accuracy with respect to the size of the grid Δx, as long as the signal is smooth (note that the major drawback is that the discrete differentiation matrices are not sparse). We set $\Delta x = 0.1$ (for $x \in [-100, 100]$), which is sufficient for the numerical errors to be undetectable. We then use the Matlab solver `ode45`, which is based on the fourth- and fifth-order Runge–Kutta–Merson method (Shampine and Reichelt 1997), with a tolerance of 10^{-8}, to solve the time-dependent problem.

Well-prepared initial data. In Fig. 2, we present a numerical simulation corresponding to the setting of Theorem 1.2; thus we compare the solution of the free-surface system with the corresponding solution of the rigid-lid system (or, more precisely, the rigid-lid approximate solution defined in Definition 3.3). We see straightforwardly that the free-surface solution closely follows the deformation of the interface and shear velocity predicted by the rigid-lid approximation, even for a relatively large value of ϱ (recall $\gamma = 0.9$ in Fig. 2c). As a matter of fact, the precision of the approximation is not predicted from Theorem 1.2: as we see from Fig. 2b, the convergence rate for ζ_2 and u_s is $\mathcal{O}(\varrho^2)$ whereas Theorem 1.2 predicts only $\mathcal{O}(\varrho)$. We can see that the main error in the rigid-lid approximation is supported on the deformation of the surface, ζ_1, as well as on the horizontal momentum, m (and more precisely the fast mode of the horizontal momentum).

Of course, such a result is predicted by Theorem 4.5 since the first-order corrector constructed in Proposition 4.2 follows precisely the preceding description. We show in Fig. 3 the precision of the improved rigid-lid approximation. We see that the main differences between the free-surface solution and the rigid-lid approximate solution have been recovered. The rate of convergence is now $\mathcal{O}(\varrho^2)$ for each variable ζ_1, ζ_2, u_s, m, in full accordance with Theorem 4.5.

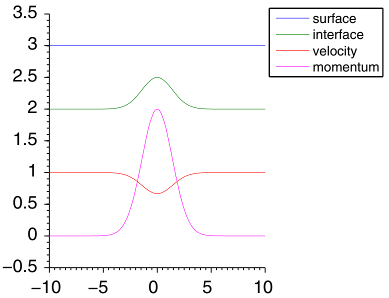
Ill-prepared initial data. We discuss now the case of ill-prepared initial data, that is, when $\zeta_1 \mid_{t=0}, m \mid_{t=0}$ are not assumed to be small. We chose to assign a nontrivial initial value only to the horizontal momentum variable m, so that the hypothesis $\alpha = \varrho$ cannot artificially modify the convergence rate (recall that the surface deviation from the flat equilibrium value is represented by $\epsilon\alpha\zeta_1$).

We plot in Fig. 4 the difference between the exact solution of the free-surface system and the approximate solution constructed in Proposition 4.6. As can be seen, there is a noticeable difference between the two solutions. Moreover, this discrepancy seems to be mainly located on the fast mode and on the variables ζ_1 and m. As a matter of fact, the variables ζ_2 and u_s present a slightly better convergence rate in Fig. 4b [around $\mathcal{O}(\varrho^{1.2})$ and $\mathcal{O}(\varrho^{1.5})$, respectively] than predicted by Proposition 4.6, namely $\mathcal{O}(\varrho)$.

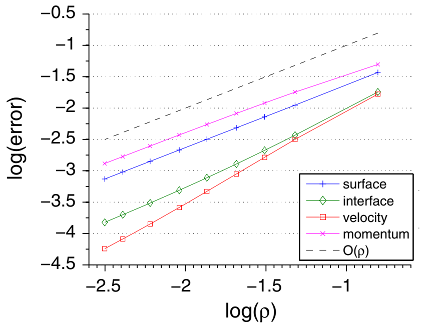
Such a result suggests the construction of a higher-order approximation, similar to the case of well-prepared initial data. Indeed, we know from Proposition 4.2 that one can construct a first-order slow-mode corrector term $(\varrho\check{\zeta}_1, 0, 0, 0)^\top$ and that its initial value plays a role in the construction of the fast-mode corrector. More precisely, one must modify the initial data of the fast-mode corrector to ensure that the full approximate solution satisfies the appropriate initial data. Using both statements of Propositions 4.2 and 4.6, we define the improved approximation for ill-prepared initial data as

$$V_{\text{app}} = V_{\text{RL}} + V^s_{\text{cor}} + V^f_{\text{cor}},$$

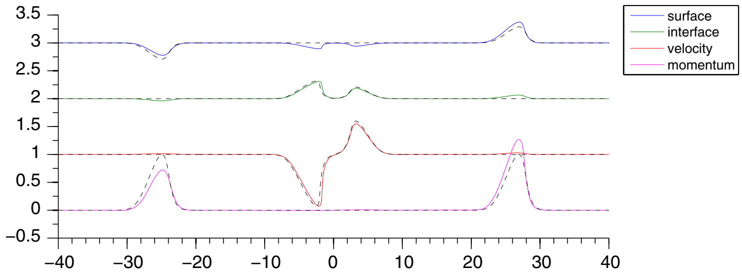
where



(a) Initial data



(b) Error with respect to ϱ (log-log scale)



(c) Flow at finite time $T = 4$

Fig. 4 Solution of free-surface system compared with approximate solution for ill-prepared initial data

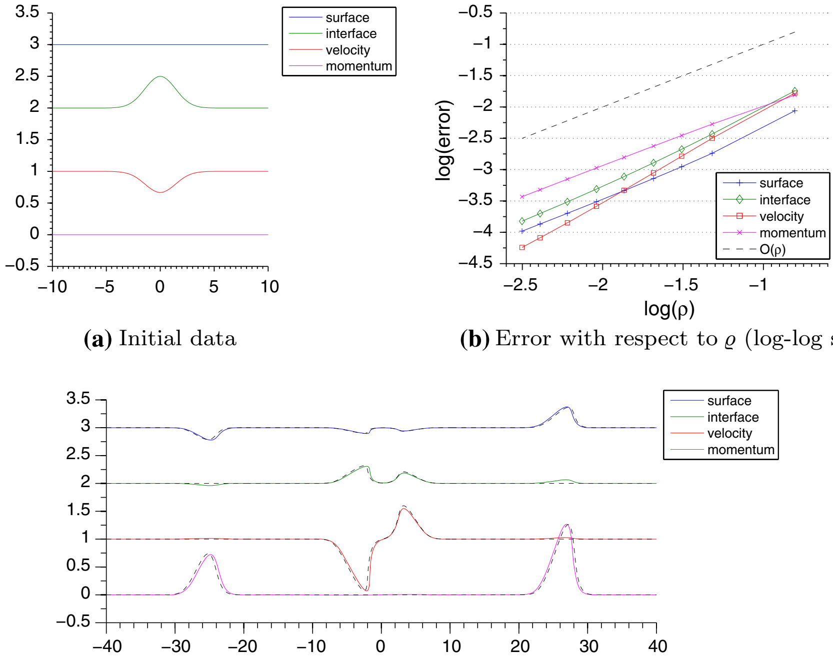
- $V_{\mathrm{RL}} \equiv (0, \eta, v, 0)^\top$ is defined by Definition 3.3;
- $V^s_{\mathrm{cor}} \equiv (\varrho \breve{\zeta}_1, 0, 0, 0)^\top$ is defined by $\breve{\zeta}_1 \equiv -\big(\eta + \frac{\delta}{2}\eta^2\big)\frac{(1-\eta)(\delta^{-1}+\eta)v^2}{(1+\delta^{-1})^2}$;
- V^f_{cor} is defined by

$$V^f_{\mathrm{cor}}(t,x) \equiv \begin{pmatrix} u_+(t,x) + u_-(t,x) \\ 0 \\ 0 \\ c(u_+(t,x) - u_-(t,x)) \end{pmatrix},$$

where $c \equiv \sqrt{1+\delta^{-1}}$ and $u_\pm$ is the unique solution to $\partial_t u_\pm \pm \frac{c}{\varrho}\partial_x u_\pm \pm \frac{3}{2c} u_\pm \partial_x u_\pm = 0$, with $u_\pm|_{t=0} = \frac{1}{2}\big(\zeta_1^0 - \varrho\breve{\zeta}_1|_{t=0} \pm c^{-1} m^0\big)$.

Note that, as was previously mentioned in Remark 4.7, this improved approximation is equivalent to the one already defined in Proposition 4.2 for well-prepared initial data. Thus this approximate solution is quite general and robust: it offers the same precision as our previously constructed approximate solutions in the well-prepared case (Proposition 4.2) as well as in the ill-prepared case (Proposition 4.6).

We investigate in Fig. 5 the accuracy of this improved approximate solution. Comparing Figs. 4c and 5c, one clearly sees that the new approximate solution shows a better resemblance than the original approximate solution; the main discrepancy seems to be corrected. However, as one can see from Fig. 5b, this apparent improvement is



(a) Initial data **(b)** Error with respect to ϱ (log-log scale)

(c) Flow at finite time $T = 4$

Fig. 5 Solution of free-surface system compared with improved approximate solution for ill-prepared initial data

not reflected in the convergence rate. Although the produced error is clearly smaller, the rate is not better than $\mathcal{O}(\varrho)$ where ζ_1 and m are involved (ζ_2 and u_s are unchanged). It is not clear to us whether a better approximate solution can be constructed, nor what explains the slightly better convergence rate on ζ_2 and u_s. Our numerical simulations indicate that there is a nontrivial coupling between the fast and slow modes at early times (when both are localized at the same place) and that the contribution of these coupling effects is of size $\approx \varrho$. Thus to take into account these coupling effects, one may have no other choice than to solve a fully coupled system, at least for a small time scale, $t = \mathcal{O}(\varrho)$.

Acknowledgments The author is grateful to Christophe Cheverry, Jean-François Coulombel, and Frédéric Rousset for helpful advice and stimulating discussions. This work was partially supported by Project ANR-13-BS01-0003-01 DYFICOLTI.

Appendix: Proof of Proposition 2.2

In this section, we detail the proof of Proposition 2.2, which follows the classical theory concerning Friedrichs-symmetrizable quasilinear systems. The proof is based on a priori energy estimates, for which the key ingredients are product and commutator estimates in Sobolev spaces. We first recall such results and refer the reader to, e.g., Alinhac and Gérard (1991), Lannes (2013) for the proof of Lemmata 5.1 and 5.3.

Lemma 5.1 *(Product estimates) Let $s \geq 0$. For all $f, g \in H^s(\mathbb{R}) \bigcap L^\infty(\mathbb{R})$, one has*

$$\big| f\, g \big|_{H^s} \lesssim \big| f \big|_{L^\infty} \big| g \big|_{H^s} + \big| f \big|_{H^s} \big| g \big|_{L^\infty}.$$

If $s \geq s_0 > 1/2$, then one deduces, thanks to a continuous embedding of Sobolev spaces,

$$\big| f\, g \big|_{H^s} \lesssim \big| f \big|_{H^s} \big| g \big|_{H^s}.$$

Let $F \in C^\infty(\mathbb{R})$ such that $F(0) = 0$. If $g \in H^s(\mathbb{R}) \bigcap L^\infty(\mathbb{R})$ with $s \geq 0$, then one has $F(g) \in H^s(\mathbb{R})$ and

$$\big| F(g) \big|_{H^s} \leq C(\big| g \big|_{L^\infty}, \big| F \big|_{C^\infty}) \big| g \big|_{H^s}.$$

Throughout the paper, we repeatedly make use of the following corollary.

Corollary 5.2 *Let $f, \zeta \in L^\infty \bigcap H^s$, with $s \geq 0$ and $h(\zeta) \equiv 1 - \zeta$, with $h(\zeta) \geq h_0 > 0$ for any $x \in \mathbb{R}$. Then one has*

$$\begin{aligned} \big|\frac{1}{h(\zeta)} f\big|_{H^s} &\leq C(h_0^{-1}, |\zeta|_{L^\infty})\big(|f|_{H^s} + |\zeta|_{H^s}|f|_{L^\infty}\big) \\ \big|f - \frac{1}{h(\zeta)} f\big|_{H^s} &\leq C(h_0^{-1}, |\zeta|_{L^\infty})\big(|\zeta|_{L^\infty}|f|_{H^s} + |\zeta|_{H^s}|f|_{L^\infty}\big). \end{aligned}$$

Proof We will use the identity

$$\frac{1}{h(\zeta)} f = \frac{1}{1-\zeta} f = f + \frac{\zeta}{1-\zeta} f.$$

By Lemma 5.1, we deduce

$$\begin{aligned} \big|\frac{1}{h(\zeta)} f\big|_{H^s} &\leq |f|_{H^s} + \big|\frac{\zeta}{1-\zeta} f\big|_{H^s} \\ &\lesssim |f|_{H^s} + \big|\frac{\zeta}{1-\zeta}\big|_{L^\infty}|f|_{H^s} + \big|\frac{\zeta}{1-\zeta}\big|_{H^s}|f|_{L^\infty}. \end{aligned}$$

The only nontrivial term to estimate is now $\big|\frac{\zeta}{1-\zeta}\big|_{H^s}$. Using that $h(\zeta) = 1 - \zeta \geq h_0 > 0$, we introduce a function $F \in C^\infty(\mathbb{R})$ such that

$$F(X) = \begin{cases} \frac{X}{1-X} & \text{if } 1 - X \geq h > 0, \\ 0 & \text{if } 1 - X \leq 0. \end{cases}$$

The function F satisfies the hypotheses of Lemma 5.1, and we have

$$\big|\frac{\zeta}{1-\zeta}\big|_{H^s} = \big|F(\zeta)\big|_{H^s} \leq C(|\zeta|_{L^\infty}, h_0^{-1})|\zeta|_{H^s}.$$

The first estimate of the lemma is proved. The second estimate is obtained in the same way using

$$f - \frac{1}{h(\zeta)} f \;=\; -\frac{\zeta}{1-\zeta} f.$$

The corollary is proved. □

The following lemma presents a generalization of the Kato–Ponce (Kato and Ponce 1988) commutator estimates due to Lannes Lannes (2006) (one has $\big|f\big|_{H^s}$ instead of $\big|\partial_x f\big|_{H^{s-1}}$ in the standard Kato–Ponce estimate).

Lemma 5.3 *(Commutator estimates) For any $s \geq 0$ and $\partial_x f, g \in L^\infty(\mathbb{R}) \bigcap H^{s-1}(\mathbb{R})$ we have*

$$\big|\, [\Lambda^s, f]g \,\big|_{L^2} \;\lesssim\; \big|\, \partial_x f \,\big|_{H^{s-1}} \big|\, g \,\big|_{L^\infty} + \big|\, \partial_x f \,\big|_{L^\infty} \big|\, g \,\big|_{H^{s-1}}.$$

Thanks to the continuous embedding of Sobolev spaces, we have for $s \geq s_0+1,\ s_0 > \frac{1}{2}$,

$$\big|\, [\Lambda^s, f]g \,\big|_{L^2} \;\lesssim\; \big|\, \partial_x f \,\big|_{H^{s-1}} \big|\, g \,\big|_{H^{s-1}}.$$

Let us now continue with the proof of Proposition 2.2. System (1.1) is quasilinear. In what follows we prove that it is Friedrichs-symmetrizable under conditions (2.2). We present below the symmetrizer of the system and compute the necessary energy estimates in Lemmata 5.5 and 5.6.

System symmetrizer. Recall that (1.1) reads $\partial_t U \;+\; A[U]\partial_x U \;=\; 0$, with

$$A[U] \;\equiv\; \begin{pmatrix} u_1 & \frac{u_2-u_1}{\varrho} & \frac{h_1}{\varrho} & \frac{h_2}{\varrho} \\ 0 & u_2 & 0 & h_2 \\ \frac{1}{\varrho} & 0 & u_1 & 0 \\ \frac{\gamma}{\varrho} & \delta+\gamma & 0 & u_2 \end{pmatrix}, \tag{5.1}$$

where we use the notation $h_1 \equiv 1 + \varrho\zeta_1 - \zeta_2$ and $h_2 \equiv \delta^{-1} + \zeta_2$. Define

$$S[U] \equiv \begin{pmatrix} \gamma & 0 & 0 & 0 \\ 0 & \gamma+\delta & 0 & u_2-u_1 \\ 0 & 0 & \gamma h_1 & 0 \\ 0 & u_2-u_1 & 0 & h_2 \end{pmatrix}. \tag{5.2}$$

We can easily check that $S[U]A[U] \equiv \Sigma[U]$ and $S[U]$ are symmetric. More precisely, we have

$$\Sigma[U] \equiv \begin{pmatrix} \gamma u_1 & \frac{\gamma(u_2-u_1)}{\varrho} & \frac{\gamma h_1}{\varrho} & \frac{\gamma h_2}{\varrho} \\ \frac{\gamma(u_2-u_1)}{\varrho} & 2(\gamma+\delta)(2u_2-u_1) & 0 & (\gamma+\delta)h_2+u_2(u_2-u_1) \\ \frac{\gamma h_1}{\varrho} & 0 & \gamma h_1 u_1 & 0 \\ \frac{\gamma h_2}{\varrho} & (\gamma+\delta)h_2+u_2(u_2-u_1) & 0 & h_2(2u_2-u_1) \end{pmatrix}. \tag{5.3}$$

We can easily check that $S[U]$ is positive definite provided that the following holds:

$$\gamma > 0 \quad ; \quad \gamma+\delta > 0 \quad ; \quad h_1 > 0 \quad ; \quad h_2 - \frac{|u_2-u_1|^2}{\gamma+\delta} > 0,$$

which is guaranteed by condition (2.2).

Energy of our system. The natural energy of our system is

$$\begin{aligned} E^s(U) &\equiv \big(S[\underline{U}]\Lambda^s U, \Lambda^s U\big) \\ &= \gamma\big|\zeta_1\big|_{H^s}^2 + (\gamma+\delta)\big|\zeta_2\big|_{H^s}^2 + \gamma\int_{\mathbb{R}} \underline{h}_1\big|\Lambda^s u_1\big|^2 \\ &\quad + \int_{\mathbb{R}} \underline{h}_2\big|\Lambda^s u_2\big|^2 + 2\int_{\mathbb{R}} (\underline{u}_2-\underline{u}_1)\big\{\Lambda^s u_2\big\}\big\{\Lambda^s \zeta_2\big\}, \end{aligned} \tag{5.4}$$

with $\underline{h}_1 \equiv 1 + \varrho\underline{\zeta}_1 - \underline{\zeta}_2$ and $\underline{h}_2 \equiv \delta^{-1} + \underline{\zeta}_2$.

In what follows, we specify the equivalence between our energy and the norm X^s offered by the well-posedness of the symmetrizer. Recall that X^s denotes the space $H^s(\mathbb{R})^4$, endowed with the following norm:

$$\big|U\big|_{X^s}^2 \; = \; \gamma\big|\zeta_1\big|_{H^s}^2 \; + \; \big|\zeta_2\big|_{H^s}^2 + \gamma\big|u_1\big|_{H^s}^2 + \big|u_2\big|_{H^s}^2.$$

Lemma 5.4 *Let $s \geq 0$ and $\underline{\zeta} \in L^\infty(\mathbb{R})$, satisfying (2.2). Then $E^s(U)$ is uniformly equivalent to the $|\cdot|_{X^s}$-norm. More precisely, there exists positive constants $C_2 = C(h_0^{-1}, \delta_{\min}^{-1}) > 0$, and $C_1 = C(\big|\underline{h}_1\big|_{L^\infty}, \big|\underline{h}_2\big|_{L^\infty}, \delta_{\max}) > 0$ such that*

$$\frac{1}{C_1}E^s(U) \; \leq \; \big|U\big|_{X^s}^2 \; \leq \; C_2 E^s(U).$$

Proof The fact that $E^s(U) \leq C_1\big|U\big|_{X^s}$ is a simple consequence of the Cauchy–Schwarz inequality, applied to (5.4), where we use that (2.2) yields $|\underline{u}_2 - \underline{u}_1|^2 < (\gamma+\delta)\underline{h}_2$.

The other inequality follows directly from (2.2). More precisely, we have

$$E^s(U) \geq \gamma \big|\zeta_1\big|_{H^s}^2 + \gamma h_0 \int_{\mathbb{R}} \big|\Lambda^s u_1\big|^2 \\ + (\gamma+\delta)\big|\zeta_2\big|_{H^s}^2 + \int_{\mathbb{R}} \underline{h}_2 \big|\Lambda^s u_2\big|^2 - 2\int_{\mathbb{R}} \sqrt{(\underline{h}_2 - h_0)(\gamma+\delta)}\big\{\Lambda^s u_2\big\}\big\{\Lambda^s \zeta_2\big\},$$

and the result is now clear. Lemma 5.4 is proved. □

We now highlight energy estimates with respect to the linearized system from (1.1), namely

$$\partial_t U \ + \ A[\underline{U}]\partial_x U \ = \ \mathcal{R}\,, \tag{5.5}$$

with given $\underline{U}$, $\mathcal{R}$.

Lemma 5.5 *(L^2 energy estimate) Set $T, M > 0$. Let $U \in L^\infty([0,T]; X^0)$ satisfy (5.5), with given $\mathcal{R} \in L^1([0,T]; X^0)$, and $\underline{U}$ satisfying (2.2), with $h_0 > 0$ (for any $t \in [0,T]$) as well as*

$$\big\|\underline{U}\big\|_{L^\infty([0,T]\times\mathbb{R})^4} + \big\|\partial_x \underline{U}\big\|_{L^\infty([0,T]\times\mathbb{R})^4} + \varrho\big\|\partial_t \underline{U}\big\|_{L^\infty([0,T]\times\mathbb{R})^4} \ \leq \ M.$$

Then there exists $C_0 \equiv C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ such that $\forall t \in [0,T]$,

$$E^0(U)(t) \leq e^{C_0 M \varrho^{-1} t} E^0(U\mid_{t=0}) + \ C_0 \int_0^t e^{C_0 M \varrho^{-1}(t-t')}\big|\mathcal{R}(t',\cdot)\big|_{X^s}\,dt'. \tag{5.6}$$

Proof Let us consider the L^2 inner product of (5.5) and $S[\underline{U}]U$:

$$\big(\partial_t U, S[\underline{U}]U\big) \ + \ \big(A[\underline{U}]\partial_x U, S[\underline{U}]U\big) \ = \ \big(\mathcal{R}, S[\underline{U}]U\big).$$

From the symmetry property of $S[\underline{U}]$, $\Sigma[\underline{U}]$, and using the definition of $E^0(U)$, we deduce

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}E^0(U) \ &= \frac{1}{2}\big(U, \big[\partial_t, S[\underline{U}]\big]U\big) - \big(\Sigma[\underline{U}]\partial_x U, U\big) + \big(\mathcal{R}, S[\underline{U}]U\big) \\ &= \frac{1}{2}\big(U, \big[\partial_t, S[\underline{U}]\big]U\big) + \frac{1}{2}\big(\big[\partial_x, \Sigma[\underline{U}]\big]U, U\big) + \big(\mathcal{R}, S[\underline{U}]U\big).\end{aligned} \tag{5.7}$$

We now estimate each of the terms on the right-hand side of (5.7).

Estimate of $\big(U, \big[\partial_t, S[\underline{U}]\big]U\big)$. We have $\big(U, \big[\partial_t, S[\underline{U}]\big]U\big) = \big(U, \mathrm{d}S[\partial_t \underline{U}]U\big)$, with

$$\mathrm{d}S[\partial_t \underline{U}] \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \partial_t(\underline{u}_2 - \underline{u}_1) \\ 0 & 0 & \gamma\partial_t(\varrho\underline{\zeta}_1 - \underline{\zeta}_2) & 0 \\ 0 & \partial_t(\underline{u}_2 - \underline{u}_1) & 0 & \partial_t\underline{\zeta}_2 \end{pmatrix}.$$

Using the Cauchy–Schwarz inequality and Lemma 5.4 we have straightforwardly

$$\big|\big(U, \big[\partial_t, S[\underline{U}]\big]U\big)\big| \le C_0 \big|\partial_t\underline{U}\big|_{L^\infty} C_2^{-1}\, \big|U\big|_{X^0}^2 \le C_0\, M\, \varrho^{-1}\, E^0(U), \tag{5.8}$$

with $C_0 = C(h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$.

Estimate of $\big(\big[\partial_x, \Sigma[\underline{U}]\big]U, U\big)$. We have $\big(\big[\partial_x, \Sigma[\underline{U}]\big]U, U\big) = \big(U, \mathrm{d}\Sigma[\underline{U}]U\big)$, with

$$\mathrm{d}\Sigma[\underline{U}] \equiv \begin{pmatrix} \gamma\partial_x\underline{u}_1 & \frac{\gamma\partial_x(\underline{u}_2 - \underline{u}_1)}{\varrho} & \frac{\gamma\partial_x(\varrho\underline{\zeta}_1 - \underline{\zeta}_2)}{\varrho} & \frac{\gamma\partial_x\underline{\zeta}_2}{\varrho} \\ \frac{\gamma\partial_x(\underline{u}_2 - \underline{u}_1)}{\varrho} & (\gamma+\delta)\partial_x(2\underline{u}_2 - \underline{u}_1) & 0 & \partial_x\big((\gamma+\delta)\underline{\zeta}_2 + \underline{u}_2(\underline{u}_2 - \underline{u}_1)\big) \\ \frac{\gamma\partial_x(\varrho\underline{\zeta}_1 - \underline{\zeta}_2)}{\varrho} & 0 & \gamma\partial_x(\underline{h}_1\underline{u}_1) & 0 \\ \frac{\gamma\partial_x\underline{\zeta}_2}{\varrho} & \partial_x\big((\gamma+\delta)\underline{\zeta}_2 + \underline{u}_2(\underline{u}_2 - \underline{u}_1)\big) & 0 & 2\partial_x\big(\underline{h}_2(2\underline{u}_2 - \underline{u}_1)\big) \end{pmatrix}.$$

As previously, the Cauchy–Schwarz inequality and Lemmata 5.1 and 5.4 yield

$$\big|\big(\Sigma[\underline{U}]\partial_x U, U\big)\big| \le C_0\, M\, \varrho^{-1}\, E^0(U), \tag{5.9}$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$.

Estimate of $\big(\mathcal{R}, S[\underline{U}]U\big)$. By the Cauchy–Schwarz inequality and Lemmata 5.1 and 5.4,

$$\big|\big(\mathcal{R}, S[\underline{U}]U\big)\big| \le C_0\, \big|U\big|_{X^s}\big|\mathcal{R}\big|_{X^s} \le C_0' E^s(U)^{1/2}\big|\mathcal{R}\big|_{X^s}, \tag{5.10}$$

with $C_0, C_0' = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$.

Estimate (5.6) is now a consequence of the Gronwall–Bihari inequality applied to the differential inequality obtained when plugging (5.8), (5.9), (5.10) into (5.7). $\square$

Lemma 5.6 *(H^s energy estimate) Set $M, T > 0$ and $s \ge s_0 + 1, s_0 > 1/2$. Let $U \in L^\infty([0,T]; X^s)$ satisfy* (5.5)*, with $\mathcal{R} \in L^1([0,T]; X^s)$, and $\underline{U} \in L^\infty([0,T]; X^s)$ satisfying* (2.2) *as well as*

$$\big\|\underline{U}\big\|_{L^\infty([0,T];X^s)} + \varrho\big\|\partial_t\underline{U}\big\|_{L^\infty([0,T];X^{s-1})} \;\le\; M.$$

Then there exists $C_0 \equiv C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ *such that, for all* $t \in [0,T]$,

$$E^s(U)(t) \le e^{C_0 M\varrho^{-1}} E^s(U\,|_{t=0}) + C_0 \int_0^t e^{C_0 M\varrho^{-1}(t-t')}\big|\mathcal{R}(t',\cdot)\big|_{X^s}\,dt'. \tag{5.11}$$

Proof As previously, we deduce from (5.5) the identity

$$\big(\Lambda^s\partial_t U, S[\underline{U}]\Lambda^s U\big) \;+\; \big(\Lambda^s A[\underline{U}]\partial_x U, S[\underline{U}]\Lambda^s U\big) \;=\; \big(\Lambda^s\mathcal{R}, S[\underline{U}]\Lambda^s U\big)\,,$$

where we recall the notation $\Lambda \equiv (\mathrm{Id} - \partial_x^2)^{1/2}$. It follows that

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}E^s(U) \;&= \frac{1}{2}\big(\Lambda^s U, \big[\partial_t, S[\underline{U}]\big]\Lambda^s U\big) - \big(S[\underline{U}]\Lambda^s A[\underline{U}]\partial_x U, \Lambda^s U\big)\\
&\quad + \;\big(\Lambda^s\mathcal{R}, S[\underline{U}]\Lambda^s U\big)\\
&= \frac{1}{2}\big(\Lambda^s U, \big[\partial_t, S[\underline{U}]\big]\Lambda^s U\big) + \frac{1}{2}\big(\big[\partial_x, \Sigma[\underline{U}]\big]\Lambda^s U, \Lambda^s U\big)\\
&\quad + \;\big(\Lambda^s\mathcal{R}, S[\underline{U}]\Lambda^s U\big)\\
&\quad - \big(S[\underline{U}]\big[\Lambda^s, A[\underline{U}]\big]\partial_x U, \Lambda^s U\big).
\end{aligned} \tag{5.12}$$

The first three terms are bounded exactly as previously when replacing U with $\Lambda^s U$. The only novelty lies in the use of continuous Sobolev embeddings, so that

$$\big\|\underline{U}\big\|_{L^\infty([0,T]\times\mathbb{R})^4} + \big\|\partial_x\underline{U}\big\|_{L^\infty([0,T]\times\mathbb{R})^4} \;\lesssim\; \big\|\underline{U}\big\|_{L^\infty([0,T];X^s)}.$$

Similarly, we have

$$\varrho\big\|\partial_t\underline{U}\big\|_{L^\infty([0,T]\times\mathbb{R})^4} \;\lesssim\; \varrho\big\|\partial_t\underline{U}\big\|_{L^\infty([0,T];X^{s-1})}.$$

The remaining term is estimated as follows. Using the commutator estimate in Lemma 5.3 we have

$$\big|\big[\Lambda^s, A[\underline{U}]\big]\partial_x U\big|_{L^2} \;\le\; C\big|\partial_x U\big|_{H^{s-1}}\big|\big[\partial_x, A[\underline{U}]\big]\big|_{H^{s-1}} \;\le\; C_0\, M\,\varrho^{-1}\,\big|U\big|_{X^s},$$

with $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$. Altogether, we deduce from (5.12)

$$\frac{1}{2}\frac{d}{dt}E^s(U) \;\le\; C_0 M\varrho^{-1}E^s(U) \;+\; C_0 E^s(U)^{1/2}\big|\mathcal{R}\big|_{X^s}.$$

Estimate (5.11) is now a consequence of the Gronwall–Bihari inequality, and the lemma is proved. ◻

Completion of Proof of Proposition 2.2 The well-posedness of system (1.1) is now a consequence of the energy estimates of Lemmata 5.5 and 5.6, following the standard strategy (we refer the reader to standard textbooks, e.g., Taylor 1997; Alinhac and Gérard 1991; Métivier 2008, for more details). More precisely, we first show that the linearized problem (5.5) is well posed, then the solution of the nonlinear problem (1.1) is obtained as the limit of an iterative scheme:

$$\partial_t U^{n+1} + A[U^n]\partial_x U^{n+1} = 0.$$

The restriction on the time scale $t \in [0, T_Q]$ is necessary to guarantee that $(U^n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and in particular that U^n is uniformly bounded with respect to n, over a time domain which can be chosen independent of n. The desired estimate on $\big|U\big|_{X^s}$ follows directly from Lemma 5.6, with $\underline{U} = U$ and $R \equiv 0$, and the corresponding estimate on $\big|\partial_t U\big|_{X^s}$ is then deduced using (1.1). The uniqueness comes from a similar estimate on the difference between two solutions, and the blow-up criterion as $t \to T_{\max}$ if $T_{\max} < \infty$ follows from standard continuation arguments. This concludes the proof of Proposition 2.2. □

References

Abgrall, R., Karni, S.: Two-layer shallow water system: a relaxation approach. SIAM J. Sci. Comput. **31**(3), 1603–1627 (2009)

Alinhac, S., Gérard, P.: Opérateurs pseudo-différentiels et théorème de Nash-Moser. Savoirs Actuels (1991)

Barros, R., Gavrilyuk, S.L., Teshukov, V.M.: Dispersive nonlinear waves in two-layer flows with free surface. I. Model derivation and general properties. Stud. Appl. Math. **119**(3), 191–211 (2007)

Benjamin, T.B.: Internal waves of finite amplitude and permanent form. J. Fluid Mech. **25**(2), 241–270 (1966)

Bona, J.L., Colin, T., Lannes, D.: Long wave approximations for water waves. Arch. Ration. Mech. Anal. **178**(3), 373–410 (2005)

Bona, J.L., Lannes, D., Saut, J.-C.: Asymptotic models for internal waves. J. Math. Pures Appl. (9) **89**(6), 538–566 (2008)

Bresch, D., Renardy, M.: Well-posedness of two-layer shallow water flow between two horizontal rigid plates. Nonlinearity **24**(4), 1081–1088 (2011)

Browning, G., Kreiss, H.-O.: Problems with different time scales for nonlinear partial differential equations. SIAM J. Appl. Math. **42**(4), 704–718 (1982)

Castro-Díaz, M.J., Fernández-Nieto, E.D., González-Vida, J.M., Parés-Madroñal, C.: Numerical treatment of the loss of hyperbolicity of the two-layer shallow-water system. J. Sci. Comput. **48**(1–3), 16–40 (2011)

Choi, W., Camassa, R.: Weakly nonlinear internal waves in a two-fluid system. J. Fluid Mech. **313**, 83–103 (1996)

Craig, W., Guyenne, P., Kalisch, H.: Hamiltonian long-wave expansions for free surfaces and interfaces. Commun. Pure Appl. Math. **58**(12), 1587–1641 (2005)

Craig, W., Guyenne, P., Sulem, C.: Coupling between internal and surface waves. Nat. Hazards **57**(3), 617–642 (2010)

de Saint-Venant, B.: Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit. C. R. Acad. Sci. Paris **73**:147–154 (1871)

Duchêne, V.: Asymptotic shallow water models for internal waves in a two-fluid system with a free surface. SIAM J. Math. Anal. **42**(5), 2229–2260 (2010)

Duchêne, V.: Boussinesq/Boussinesq systems for internal waves with a free surface, and the KdV approximation. M2AN. Math. Model. Numer. Anal. **46**, 145–185 (2011)

Duchêne, V.: Decoupled and unidirectional asymptotic models for the propagation of internal waves. M3AS. Math. Models Methods Appl. Sci. **24**(01) (2014)

Gill, A.E.: Atmosphere-Ocean Dynamics. International Geophysics Series, vol. 30. Academic Press (1982)

Grimshaw, R., Pelinovsky, E., Poloukhina, O.: Higher-order korteweg-de vries models for internal solitary waves in a stratified shear flow with a free surface. Nonlinear Processes Geophys. **9**, 221–235 (2002)

Guyenne, P., Lannes, D., Saut, J.-C.: Well-posedness of the Cauchy problem for models of large amplitude internal waves. Nonlinearity **23**(2), 237–275 (2010)

Helfrich, K.R., Melville, W.K.: Long nonlinear internal waves. Annu. Rev. Fluid Mech. **38**, 395–425 (2006)

Jackson, C.R.: An atlas of internal solitary-like waves and their properties (2004). Accessible at url URL http://www.internalwaveatlas.com/Atlas2_index.html

Kato, T.: The Cauchy problem for quasi-linear symmetric hyperbolic systems. Arch. Ration. Mech. Anal. **58**(3), 181–205 (1975)

Kato, T.: Perturbation theory for linear operators. Classics in Mathematics. Springer, Berlin (1995). Reprint of the 1980 edition

Kato, T., Ponce, G.: Commutator estimates and the Euler and Navier-Stokes equations. Commun. Pure Appl. Math. **41**(7), 891–907 (1988)

Klainerman, S., Majda, A.: Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids. Commun. Pure Appl. Math. **34**(4), 481–524 (1981)

Lannes, D.: Secular growth estimates for hyperbolic systems. J. Differ. Equ. **190**(2), 466–503 (2003)

Lannes, D.: Sharp estimates for pseudo-differential operators with symbols of limited smoothness and commutators. J. Funct. Anal. **232**(2), 495–539 (2006)

Lannes, D.: A stability criterion for two-fluid interfaces and applications. Arch. Ration. Mech. Anal. **208**(2), 481–567 (2013)

Lannes, D.: The water waves problem. Mathematical Surveys and Monographs, vol. 188. American Mathematical Society (2013)

Leonardi, D.: Internal and Surface Waves in a Two-Layer Fluid. PhD thesis, University of Illinois (2011)

Long, R.R.: On the Boussinesq approximation and its role in the theory of internal waves. Tellus **17**(1), 46–52 (1965)

Métivier, G.: Para-differential calculus and applications to the Cauchy problem for nonlinear systems. Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series, vol. 5 (2008)

Schneider, G., Wayne, C.E.: The long-wave limit for the water wave problem. I. The case of zero surface tension. Commun. Pure Appl. Math. **53**(12), 1475–1535 (2000)

Shampine, L.F., Reichelt, M.W.: The MATLAB ODE suite. SIAM J. Sci. Comput. **18**(1), 1–22 (1997)

Stewart, A.L., Dellar, P.J.: Multilayer shallow water equations with complete coriolis force. Part 3. Hyperbolicity and stability under shear. J. Fluid Mech. **723**, 289–317, 5 (2013)

Taylor, M.E.: Partial differential equations. III Nonlinear equations. Applied Mathematical Sciences, vol. 117. Springer (1997)

Trefethen, L.N.: Spectral methods in MATLAB. Software, Environments, and Tools. Society for Industrial and Applied Mathematics (SIAM), vol. 10 (2000).