



**HAL**  
open science

# On the rigid-lid approximation for two shallow layers of immiscible fluids with small density contrast

Vincent Duchêne

► **To cite this version:**

Vincent Duchêne. On the rigid-lid approximation for two shallow layers of immiscible fluids with small density contrast. 2013. hal-00861093v2

**HAL Id: hal-00861093**

**<https://hal.science/hal-00861093v2>**

Preprint submitted on 9 Oct 2013 (v2), last revised 4 Feb 2014 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the rigid-lid approximation for two shallow layers of immiscible fluids with small density contrast

Vincent Duchêne

October 10, 2013

## 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. 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 small density contrast, the flow may be accurately described as a 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 non-small initial barotropic mode.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Presentation of the models, and main result . . . . .	2
<b>2</b>	<b>Preliminary results</b>	<b>6</b>
<b>3</b>	<b>Proof of the main result</b>	<b>8</b>
3.1	Construction of the approximate solution . . . . .	10
3.2	Properties of the system and its symmetrizer . . . . .	12
3.3	Completion of the proof . . . . .	14
<b>4</b>	<b>Decomposition of the flow</b>	<b>17</b>
4.1	Improved approximate solution . . . . .	18
4.2	The case of ill-prepared initial data . . . . .	23
4.3	Discussion and numerical simulations . . . . .	28
<b>A</b>	<b>Proof of Proposition 2.2</b>	<b>32</b>

## 1 Introduction

### 1.1 Motivation

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

(see, *e.g.*, [15, 18, 19] and references therein). The study of these internal waves has attracted a considerable amount of attention in the past decades, and lead to a vast collection of various models. In order to simplify the setting, two approximations are commonly used in the literature, namely 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. Acknowledgedly, these two assumptions are related: a fixed amount of energy generates a much smaller displacement on the air–water interface than on the fresh-salted water interface, because the ratio of mass densities across the interface is negligible in the former case when compared to the latter.

The ambition of this article is to offer a rigorous justification of the above presumption. We restrict ourselves to one of the simplest possible setting, that is two infinite, two-dimensional layers of immiscible fluids with constant density, above a flat bottom. Moreover, we consider extremely shallow layers, so that the horizontal velocity field is assumed to be constant throughout the depth of each layer; thus we study the so-called Saint-Venant [11], or shallow-water systems. Even in that much simplified setting, we will come across serious difficulties, which come 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 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.* non-vanishing with the density contrast) is challenging.

To our knowledge, very few works are concerned with the validity of the aforementioned approximations, despite the early concerns expressed by Long [27] and Benjamin [4]. Grimshaw, Pelinovsky, Poloukhina [16], Craig, Guyenne, Kalisch [9], Craig, Guyenne, Sulem [10] and the author [13] derived and compared 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 maybe 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 [26] studied in much details 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 vanishing density contrast.

## 1.2 Presentation of the models, and main result

In this section, we present the two models we study, namely the shallow-water (or Saint-Venant) systems in the free-surface and rigid-lid configuration; see Figure 1. We briefly describe 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.

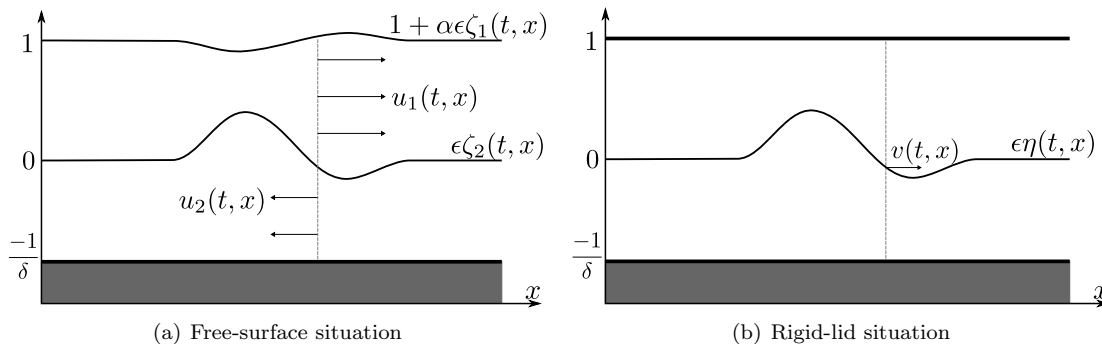


Figure 1: Sketch of the domain in the two different situations

**The free-surface system.** Let us first introduce the shallow-water model with free surface, that we simply refer to as *free-surface system*.

$$\begin{cases} \alpha \partial_t \zeta_1 + \partial_x(h_1 u_1) + \partial_x(h_2 u_2) = 0, \\ \partial_t \zeta_2 + \partial_x(h_2 u_2) = 0, \\ \partial_t u_1 + \alpha \frac{\delta + \gamma}{1 - \gamma} \partial_x \zeta_1 + \frac{\epsilon}{2} \partial_x(|u_1|^2) = 0, \\ \partial_t u_2 + (\delta + \gamma) \partial_x \zeta_2 + \gamma \alpha \frac{\delta + \gamma}{1 - \gamma} \partial_x \zeta_1 + \frac{\epsilon}{2} \partial_x(|u_2|^2) = 0, \end{cases} \quad (1.1)$$

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

This system has been obtained<sup>1</sup> in [8,9], and rigorously justified in [12] as an asymptotic model for a system of two layers of immiscible, homogeneous, ideal, incompressible fluid under the only influence of gravity (the so-called full Euler system). It describes the evolution of the deformation of the surface,  $\zeta_1$ , the interface,  $\zeta_2$ , and the velocity of the fluid in the upper (resp. lower) layer,  $u_1$  (resp.  $u_2$ ).<sup>2</sup> 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 the functions  $\zeta_1(t, x)$ ,  $\zeta_2(t, x)$  (see Figure 1(a)).

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

$\delta$  represents the ratio of the upper-layer to the lower-layer depth;

$\gamma$  represents the ratio of the mass density between the two fluids;

$\epsilon$  represents the maximal deformation of the interface, divided by the upper-layer depth;

$\alpha$  represents the ratio of the maximal deformation of the surface to the one 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  $\mu$  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  $\mathcal{O}(\mu)$  are neglected.*

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 [3], after manipulating the equations, one may obtain:

- Conservation of horizontal momentum:

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

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:

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

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

<sup>1</sup>The models presented in these works are not limited to 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 non-dimensionalizing step. We chose our scaling in order to set the typical velocity of the internal wave (obtained by solving explicitly the linear system, setting  $\alpha = \epsilon = 0$ ) as  $c_0 = \pm 1$ , consistently with the rigid-lid system (1.2).

<sup>2</sup>The Saint-Venant model is constructed from the assumption that the horizontal velocity field may be accurately described as constant throughout the depth of each layer of fluid.

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

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

Here,  $\eta$  represents the deformation of the interface, and  $v$  the shear velocity (typically,  $v = u_2 - \gamma u_1$ ; see below and Figure 1(b)). Again,  $h_1, h_2$  denote the depth of the upper (resp. lower) layers, thus  $h_1 = 1 - \epsilon \eta$  and  $h_2 = 1/\delta + \epsilon \eta$ . Parameters  $\gamma, \delta, \epsilon$  are defined as previously.

This model has been rigorously justified in [6], 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

$$\partial_x(h_1 u_1) + \partial_x(h_2 u_2) = 0.$$

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  $v \equiv u_2 - \gamma u_1$ , one obtains

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

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 at giving a rigorous confirmation of the above 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*

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

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

$$|\zeta_2^0|_{H^{s+1}} + |u_2^0 - \gamma u_1^0|_{H^{s+1}} \leq M \quad \text{and} \quad \frac{\alpha}{\varrho} |\zeta_1^0|_{H^{s+1}} + |\gamma h_1 u_1^0 + h_2 u_2^0|_{H^{s+1}} \leq M \varrho \quad (1.3)$$

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

$$\forall x \in \mathbb{R}, \quad h_1^0(x) \geq h_0 > 0 \quad ; \quad h_2^0(x) - \epsilon^2 \frac{|u_2^0(x) - u_1^0(x)|^2}{\gamma + \delta} \geq h_0 > 0, \quad (1.4)$$

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

*Then there exists positive constants  $T^{-1}, C$ , depending only on  $M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}, \frac{1}{s_0 - 1/2}$ , such that the following holds.*

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  $0 \leq t < T/\max\{\epsilon M, \varrho\}$ ,*

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

*and*

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

**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 do not reveal any limitation that would appear in the weakly non-linear case,  $\epsilon M = \mathcal{O}(\varrho)$ . On the contrary, we know that in the latter case (see Proposition 2.2 and Remark 2.3), the system (1.1) is well-posed over time  $\mathcal{O}((\epsilon M)^{-1})$ , without the additional condition (1.3). Moreover, it would not be difficult to obtain an asymptotic description of the solution similar to the one obtained by the author in [13] (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) complies to similar description (with only two counter-propagating waves), thus the two solutions are easily compared. One ambition of the present article is to extend such results to stronger nonlinearities; see in particular Theorem 4.5 and Proposition 4.7.*

*In order to acknowledge the fact that we are interested in strong nonlinearities, and to ease the reading, we set  $\epsilon \equiv 1$  in the following, without loss of generality.*

**Remark 1.4.** *The factor  $\frac{\alpha}{\varrho}$  in front of  $\zeta_1$  is natural in our context. Indeed, one easily deduces from the aforementioned conservation of energy for (1.1) that*

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

*so that without any further assumption than a finite initial energy, we know that  $\gamma^{1/2} \frac{\alpha}{\varrho} |\zeta_1|_{L^2}$  remains bounded as long as the solution is well-defined. We set  $\alpha \equiv \varrho$  in the following, without loss of generality.*

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

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

with fixed  $0 < \delta_{\min} \leq \delta_{\max} < \infty$ . The interesting limit is therefore  $\varrho \rightarrow 0$  or, equivalently,  $\gamma \rightarrow 1$ . Except for Section 2 and Appendix A, we additionally impose  $0 < \gamma_{\min} \leq \gamma$ , with  $\gamma_{\min}$  fixed.

**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) has been studied in details in [17] (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 obtains easily the well-posedness of the Cauchy problem for the free-surface equation with standard energy methods; we state the result in Proposition 2.2, and postpone its proof to Appendix A. However, the resulting time of existence is only of size  $T \gtrsim \varrho$ . One purpose of our work is to obtain a control of the energy over large time (*i.e.* uniform with respect to  $\varrho$  small), as well as describing the asymptotic behavior of the solution when  $\varrho$  vanishes.

Let us mention 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 Section 4, we discuss several natural developments around Theorem 1.2, namely

- The construction of a first-order corrector term in order to reach a higher precision. In particular, we describe the asymptotic behavior of the small deformation at the surface.
- The case of ill-prepared initial data, that is data failing to meet the smallness assumption in (1.3).

On both counts, the relevant notion lies in a decomposition between fast mode and slow mode (or barotropic and baroclinic mode), that we precise therein. Finally, Section 4.3 also contains a discussion on the different 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 depend non-decreasingly on  $\delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}$ , and often on  $\frac{1}{s_0-1/2}$ , such dependency being non-necessarily 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  $\frac{a}{b} \rightarrow 1$  ( $\varrho \rightarrow 0$ ).

We denote by  $C(\lambda_1, \lambda_2, \dots)$  a nonnegative constant depending on the parameters  $\lambda_1, \lambda_2, \dots$ , and whose dependence on the  $\lambda_j$  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

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

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

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

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  $\Lambda$  is the fractional derivative  $\Lambda = (\operatorname{Id} - \partial_x^2)^{1/2}$ .

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:

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

Except in Section 2 and Appendix A, we assume that  $\gamma$  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 as well as the Sobolev norms with respect to the spatial variable  $x$ , with  $(u, v) = (u(t, \cdot), v(t, \cdot))$ ,  $\|u\|_{L^2} = \|u(t, \cdot)\|_{L^2}$ , and  $\|u\|_{H^s} = \|u(t, \cdot)\|_{H^s}$ , respectively.

For  $T > 0$  and  $X$  a functional space, we denote  $L^\infty([0, T); X)$ , 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:

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

Finally,  $C^k([0, T); X)$  denote 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  $U^0 = (\zeta^0, v^0)^\top \in H^s(\mathbb{R})^2$  be such that there exists  $h_0 > 0$  with

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

There exists  $T_{\max} > 0$  and a unique  $U_{\text{RL}} = (\eta, v)^\top \in C([0, T_{\max}); H^s(\mathbb{R})^2) \cap C^1([0, T_{\max}); H^{s-1}(\mathbb{R})^2)$ , maximal solution to (1.2) (with  $\epsilon = 1$ ), with initial data  $U_{\text{RL}}|_{t=0} = U^0$ .

Moreover, there exists constants  $0 < C_0, T^{-1} \leq |U^0|_{H^s(\mathbb{R})^2} C(|U^0|_{H^s(\mathbb{R})^2}, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$  such that one has  $T_{\max} \geq T$ , and for any  $t \in [0, T]$ ,

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

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

This result has been precisely expressed in [17, Theorem 1], and follows from standard techniques on quasilinear, Friedrichs-symmetrizable systems. More precisely, the existence and uniqueness of a solution follows from energy estimates on the linearized equation, of which the estimate above is a

particular case. In order 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 in Proposition 2.2, below, are also presented in [17, Corollary 1].

Let us now turn to the free-surface system, (1.1). We 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-by-4 matrix, and  $A_1(U)$  a linear mapping into 4-by-4 matrices.

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

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

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, thus lowering the *a priori* time of existence. We state in Proposition 2.2, below, the well-posedness of the Cauchy problem as given by standard energy methods on quasilinear, Friedrichs-symmetrizable systems; remark that the time of existence of the solution is restricted to the poor  $T_{\max} \gtrsim \varrho$ . This timescale is intuitively seen from a change of variable: one can set  $U(t, \cdot) \equiv \tilde{U}(t/\varrho, \cdot)$ , where  $\tilde{U}$  satisfies

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

and  $\varrho A[U] \equiv \varrho A_0 + \varrho A_1(U)$ , with the matrix  $\varrho A_0$  and the linear mapping  $\varrho A_1$  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  $U^0 \equiv (\zeta_1^0, \zeta_2^0, u_1^0, u_2^0)^\top \in X^s$  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 exists positive constants  $0 < C_0, T^{-1} \leq |U^0|_{X^s} C(|U^0|_{X^s}, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ , such that one has  $T_{\max} > T\varrho$ ,  $U(t, \cdot)$  satisfies (2.2) for any  $t \in [0, T\varrho]$  (with  $h_0/2$  replacing  $h_0$ ), and

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

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

- $|U|_{L^\infty([0, t] \times \mathbb{R})^4}$  or  $|\partial_x U|_{L^\infty([0, t] \times \mathbb{R})^4}$  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 Appendix A, so as not to interrupt the flow of the text.



**Remark 2.3.** Notice that the uniform time of existence,  $T \approx 1$ , is recovered for sufficiently small initial data:  $|U_0|_{X^s} = \mathcal{O}(\varrho)$ . This result can be viewed through the following change of unknown:  $U \equiv \varrho \check{U}$ . The function  $\check{U}$  satisfies

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

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

### 3 Proof of the 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 non-linear contribution is uniformly bounded with respect to  $\varrho$ . In order to do so, we shall use different variables. Considering the conservation of horizontal momentum displayed in Section 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:

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

if and only if

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

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

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

We still refer to this system as 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  $T > 0$ , given. Assume that for any  $t \in [0, T]$ , one has

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

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 above non-vanishing 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  $V \in C([0, T]; H^s(\mathbb{R})^4)$  (resp.  $U \in C([0, T]; H^s(\mathbb{R})^4)$ ) is deduced from the corresponding control of  $U$  (resp.  $V$ ), using product estimates in Lemma A.1, as well as Corollary A.2. The fact that  $U$  satisfies (1.1) if  $V$  satisfies (3.3), and conversely, demands somewhat tedious but straightforward computations, that we leave to the reader.  $\square$

**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 later on 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 subsequent subsection. More importantly, one sees that the only terms factored by  $\varrho^{-1}$  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 we continue with the detailed analysis in the following subsections. We first introduce some notations, used thereafter. We rewrite the hyperbolic system (3.3) as

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

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, obtained by setting  $\epsilon = 0$  in (1.1);
- $B[\cdot]$  contains the nonlinear contribution: it is uniformly bounded with respect to  $\varrho$ .

In Section 3.1, we construct an approximate solution,  $V_{\text{app}}$ , 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

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

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 (*i.e.* bounded from below uniformly with respect to  $\varrho$ ), and Theorem 1.2 quickly follows (see Section 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 has to 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 in order to deduce a uniform control of  $\partial_t W$  from the corresponding control of  $\partial_x W$ , as once again this would yield a destructive  $\mathcal{O}(\varrho^{-1})$  factor.

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

Issue (ii) asks for 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  $L_{(0)} = \lim_{\gamma \rightarrow 1} L_\varrho$ ; see below. It follows that  $|\Pi \partial_t W|_{X^{s-1}} \lesssim |W|_{X^s}$ , uniformly with respect  $\varrho$  small. As for the other component, one shows that  $T[\cdot](\text{Id} - \Pi) = T_0 + \mathcal{O}(\varrho)$  with  $T_0$  a constant matrix, so a factor of size  $\mathcal{O}(\varrho)$  is gained after differentiation or commutation with this operator.

The detailed energy estimates are computed in Section 3.3.

There is an intuitive explanation for the reason why the above claims hold. By precisely analyzing the 4-by-4 matrix  $L_\varrho$ :

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

one may check that for  $\varrho$  sufficiently small,  $L_\varrho$  has four distinct, real eigenvalues, namely

$$\lambda_\pm^f(\varrho) = \pm \sqrt{1 + \delta^{-1}} + \mathcal{O}(\varrho^2) \quad ; \quad \lambda_\pm^s(\varrho) = \pm \varrho + \mathcal{O}(\varrho^3).$$

The linear theory thus predicts that the flow can be decomposed as the superposition of four waves, propagating at velocity  $c_{\pm}^f \sim \pm \frac{\sqrt{1+\delta^{-1}}}{\varrho}$ , and  $c_{\pm}^s \sim \pm 1$ , which we name *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 denote  $L_{(0)} \equiv \lim_{\gamma \rightarrow 1} L_{\varrho} \equiv \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1+\delta^{-1} & 0 & 0 & 0 \end{pmatrix}$ , then one easily checks that the eigenvectors corresponding to the two non-zero eigenvalues of  $L_{(0)}$  are orthogonal to the kernel of  $L_{(0)}$ . Therefore, roughly speaking, the slow mode is supported by variables  $\zeta_2$  and  $u_s$ , while the fast mode is supported by variables  $\zeta_1$  and  $m$ . We take advantage of this fact by treating separately the slow mode terms (multiplying by  $\Pi$ , the orthogonal projector onto the kernel of  $L_{(0)}$ ) and fast mode terms (multiplying by  $\text{Id} - \Pi$ , the orthogonal projector onto the space spanned by the other eigenvectors of  $L_{(0)}$ ). The former contributions are easily controlled as time derivation do not induce destructive  $\mathcal{O}(\varrho^{-1})$  factor. As for the latter, the property  $T[\cdot](\text{Id} - \Pi) = T_0 + \mathcal{O}(\varrho)$  reflects the fact that the corresponding eigenvalues are well separated; thus the perturbation by  $\varrho B[\cdot]$  typically yield nonlinear deviations of size  $\mathcal{O}(\varrho)$ , following standard perturbation theory [21]. Finally, the desired property on  $\Upsilon[\cdot]$  is easily checked:

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

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

### 3.1 Construction of the 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.

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

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

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

$$\forall t \in [0, T], \quad |V_{\text{RL}}|_{X^{s+1}} + |\partial_t V_{\text{RL}}|_{X^s} \leq C_1. \quad (3.6)$$

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

$$\forall t \in [0, T], \quad |V_{\text{rem}}|_{X^{s+1}} + |\partial_t V_{\text{rem}}|_{X^s} \leq C_2 \varrho, \quad (3.7)$$

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

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

**Remark 3.5.** *The explicit formula for  $V_{\text{rem}}$ , which is precisely displayed in the proof, below, 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 Section 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  $C_1, T^{-1} \leq MC(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$  such that  $V_{\text{RL}} \in C([0, T]; X^{s+1})$  is well-defined by Definition 3.3, and (3.6) holds.

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

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

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

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

$$|\check{\zeta}_1|_{H^{s+1}} + |\check{m}|_{H^{s+1}} + |\partial_t \check{\zeta}_1|_{H^s} + |\partial_t \check{m}|_{H^s} \leq C_2, \quad (3.10)$$

and

$$|r_1|_{H^s} + |r_2|_{H^s} + |r_3|_{H^s} + |r_4|_{H^s} \leq C_3 \varrho (M + \varrho). \quad (3.11)$$

In order to ease the reading of the argument, we first assume that (3.10) holds, and see how  $\check{\zeta}_1, \check{m}$  can be naturally chosen so that (3.11) is satisfied. Our choice for  $\check{\zeta}_1, \check{m}$  is precisely stated in (3.14) and (3.16), below, 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

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

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

Let us recall that  $V_{\text{RL}}$  satisfies (3.6), and also (2.1). Thus one can apply the product estimates in Lemma A.1 as well as Corollary A.2 (we also recall that by definition,  $1 - \gamma = \varrho^2(\gamma + \delta)$ ), to deduce

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

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

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

$$r_3 = \frac{1}{2} \partial_x \left( \left\{ \frac{\gamma h_1^2 - h_2^2}{\gamma(h_1 + h_2)^2} - \frac{h_1^2 - \gamma h_2^2}{(h_1 + \gamma h_2)^2} \right\} v^2 + \frac{\gamma(\varrho^2 \check{m} + h_1 v)^2 - \gamma^2 h_1^2 v^2 + h_2^2 v^2 - (\varrho^2 \check{m} - h_2 v)^2}{\gamma(h_1 + h_2)^2} \right),$$

so that one has as above,

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

Let us now look at the fourth equation in (3.9). Note that one has

$$\begin{aligned} & \gamma(h_1 + h_2) \partial_x \check{\zeta}_1 + (\gamma + \delta) h_2 \partial_x \eta + \partial_x \left( \frac{h_1 h_2 (h_1 + \gamma h_2) v^2}{\gamma(h_1 + h_2)^2} \right) \\ &= \partial_x \left( \gamma((1 + \delta^{-1}) \check{\zeta}_1 + \frac{\varrho^2}{2} \check{\zeta}_1^2) + (\gamma + \delta)(\delta^{-1} \eta + \frac{1}{2} \eta^2) + \frac{h_1 h_2 (h_1 + \gamma h_2) v^2}{\gamma(h_1 + h_2)^2} \right). \end{aligned}$$

It is now clear that one can choose

$$\check{\zeta}_1 \equiv -\left(\eta + \frac{\delta}{2} \eta^2\right) - \frac{(1 - \eta)(\delta^{-1} + \eta)v^2}{\gamma(1 + \delta^{-1})^2}, \quad (3.14)$$

so that the above is of size  $\mathcal{O}(\varrho^2)$ . More precisely, and using once again (3.10), one has

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

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

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

We now recall that  $\eta, v$  satisfies (1.2), so that one deduces explicitly  $\partial_t \check{\zeta}_1$  from (3.14), and

$$\left| \partial_t \check{\zeta}_1 - \partial_x \left( \frac{h_1 h_2 v}{h_1 + \gamma h_2} \right) \right|_{H^s} \leq M^2 C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}).$$

Now, one can check that by choosing

$$\check{m} \equiv \frac{\delta}{1+\delta} v, \quad (3.16)$$

it follows

$$\left| \frac{h_1 h_2 v}{h_1 + \gamma h_2} + \check{m} - \frac{\gamma+\delta}{\gamma} \frac{h_1 h_2 v}{h_1 + h_2} \right|_{H^s} \leq (M^2 + M\varrho^2) C(M, h_0^{-1}, C_2, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}),$$

so that estimates (3.6) and (3.10) yield

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

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  $V_{\text{RL}}$  in (3.6), the corresponding estimate on  $V_{\text{rem}} \equiv (\varrho \check{\zeta}_1, 0, 0, \varrho^2 \check{m})^\top$ : (3.10), or equivalently (3.7). Proposition 3.4 is proved.  $\square$

### 3.2 Properties of the system and its symmetrizer

This section is dedicated to preliminary results on the new free-surface system (3.3) and its symmetrizer, which allow the energy analysis of the subsequent subsection.

We recall here that (3.3) has been 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  $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top$  solution to (3.3), through the change of variable (3.1); see Lemma 3.1. In other words, we have an explicit

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

(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 non-vanishing depth condition is satisfied:

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

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

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

one may rewrite (3.3) (after multiplication with the appropriate operator)

$$dF[V] \partial_t V + A[F(V)] dF[V] \partial_x V = 0,$$

where  $dF[V]$  is the Jacobian matrix of  $F$ . In other words, recalling earlier notation in (3.4), one has

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

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

**Lemma 3.6.** *Let  $S[\cdot]$  be a symmetrizer of (1.1); e.g. (A.2). Then  $T[\cdot] \equiv (\mathrm{d}F[\cdot])^\top S[F(\cdot)] \mathrm{d}F[\cdot]$  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

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

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

$$(\mathrm{d}F[\cdot])^\top S[\cdot] \mathrm{d}F[\cdot] (\mathrm{d}F[V])^{-1} A[F(V)] \mathrm{d}F[V] = (\mathrm{d}F[\cdot])^\top \Sigma[F(V)] \mathrm{d}F[V]$$

is symmetric, and this concludes the proof.  $\square$

We conclude that one can construct an explicit symmetrizer of system (3.3), using  $S[\cdot]$  given in (A.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 at stake, which are easily checked with the use of a computer algebra system, such as Maple.

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

$$\|B[V]\|_X \leq C_0 |V|_X, \quad \|B[V] - B[W]\|_X \leq C_0 |V - W|_X, \quad (3.19)$$

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

*Proof.* Let us recall that  $B[\cdot]$  is complicated, but 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 A.1 and Corollary A.2 (since (3.18) holds), and the result easily follows.  $\square$

**Lemma 3.8.** *Denote  $T[\cdot] \equiv (\mathrm{d}F[\cdot])^\top S[\cdot] \mathrm{d}F[\cdot]$  and  $\Upsilon[V] \equiv (\mathrm{d}F[\cdot])^\top \Sigma[F(V)] \mathrm{d}F[V]$ , with  $S[\cdot]$  and  $\Sigma[\cdot]$  defined in (A.2), (A.3). Let  $V \in X$  such that  $F(V)$  satisfies (2.2) with  $h_0 > 0$ . Then there exists  $C_0 = C(|V|_X, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$  such that one has*

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

$$\frac{1}{C_0} |W|_{L^2}^2 \leq (T[V]W, W) \leq C_0 |W|_{L^2}^2. \quad (3.20)$$

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

$$\|T[V]\|_X \leq C_0 \quad ; \quad \|\Upsilon[V]\|_X \leq \varrho^{-1} C_0. \quad (3.21)$$

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

$$\|\partial_\varkappa(T[V])\|_X \leq C_0 |\partial_\varkappa V|_X \quad ; \quad \|\partial_\varkappa(\Upsilon[V])\|_X \leq C_0 |\partial_\varkappa V|_X \quad (3.22)$$

and

$$\|\partial_\varkappa(T[V])(\mathrm{Id} - \Pi)\|_X \leq \varrho C_0 |\partial_\varkappa V|_X, \quad (3.23)$$

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 has been already stated in Lemma 3.6. Estimate (3.20) follows from Lemma A.5, noticing that  $\gamma \geq \gamma_{\min} > 0$  ensures that the  $L^2(\mathbb{R})^4$ -norm is equivalent to the  $X^0$ -norm.

Estimate (3.21) is straightforwardly deduced from the corresponding one in Lemma 3.7, and similar control of  $\|\mathrm{d}F[V]\|$ . The first estimate in (3.22) is obtained similarly.

Finally, the second estimate in (3.22) as well as (3.23) can be checked with the help of a computer algebra system (we have to ensure that first order terms in  $\varrho$  are all constant). Once again 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 A.1 and Corollary A.2.  $\square$

### 3.3 Completion of the proof

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

$$\partial_t W + \frac{1}{\varrho} (L_\varrho + \varrho B[V]) \partial_x W = \mathcal{R}, \quad (3.24)$$

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

The following Lemma presents an *a priori* energy estimate on  $W$  satisfying the above system, from which our desired result is based on.

**Lemma 3.9.** *Let  $s \geq s_0 + 1$ ,  $s_0 > 1/2$ , and  $W$  a strong solution to (3.24), with  $W|_{t=0} \in X^s$ . Assume that there exists  $M, T, h_0 > 0$  such that  $F(V)$  satisfies (2.2) and*

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

Then one has

$$\forall t \in [0, T], \quad |W(t, \cdot)|_{X^s} \leq C_0 |W(0, \cdot)|_{X^s} e^{C_0 M t} + C_0 \int_0^t e^{C_0 M (t-t')} |\mathcal{R}(t', \cdot)|_{X^s} dt'. \quad (3.25)$$

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

*Proof.* We compute the inner product of (3.24) with  $T[V] \Lambda^{2s} W$ , and obtain

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

where  $T[\cdot]$  and  $\Upsilon[\cdot]$  have been defined in the previous subsection.

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

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} E^s(W) &= \frac{1}{2} ([\partial_t, T[V]] \Lambda^s W, \Lambda^s W) + \frac{1}{2} ([\partial_x, \Upsilon[V]] \Lambda^s W, \Lambda^s W) \\ &\quad - ([\Lambda^s, T[V]] \partial_t W, \Lambda^s W) - ([\Lambda^s, \Upsilon[V]] \partial_x W, \Lambda^s W) + (\Lambda^s T[V] \mathcal{R}, \Lambda^s W), \end{aligned} \quad (3.26)$$

where we define

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

We estimate below each of the terms in the right-hand side of (3.26).

*Estimate of  $([\partial_t, T[V]] \Lambda^s W, \Lambda^s W)$ .* From (3.22) in Lemma 3.8 (with  $X = L^\infty(\mathbb{R}^4)$ ), one has

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

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

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

It follows from the above and Cauchy-Schwarz inequality that

$$|([\partial_t, T[V]] \Lambda^s W, \Lambda^s W)| \leq C_0 M |W|_{X^s}^2, \quad (3.27)$$

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

*Estimate of  $([\partial_x, \Upsilon[V]] \Lambda^s W, \Lambda^s W)$ .* As above, Cauchy-Schwarz inequality and Lemma 3.8 yield

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

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

$$|([\partial_x, \Upsilon[V]] \Lambda^s W, \Lambda^s W)| \leq C_0 M |W|_{X^s}^2, \quad (3.28)$$

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

*Estimate of  $(\Lambda^s T[V] \mathcal{R}, \Lambda^s W)$ .* We apply Cauchy-Schwarz inequality and (3.21) in Lemma 3.8. One deduces

$$(\Lambda^s T[V] \mathcal{R}, \Lambda^s W) \leq C_0 |W|_{X^s} |\mathcal{R}|_{X^s}, \quad (3.29)$$

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

*Estimate of  $([\Lambda^s, \Upsilon[V]] \partial_x W, \Lambda^s W)$ .* We make use of Kato-Ponce's commutator estimate recalled in Lemma A.3. It follows

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

From (3.22) in Lemma 3.8, and since  $X^{s-1}$  is a Banach algebra, one has

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

It follows

$$|([\Lambda^s, \Upsilon[V]] \partial_x W, \Lambda^s W)| \leq C_0 M |W|_{X^s}^2, \quad (3.30)$$

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

*Estimate of  $([\Lambda^s, T[V]] \partial_t W, \Lambda^s W)$ .* As above, Kato-Ponce's commutator estimate yields

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

Unfortunately, making use of the identity (3.24) only yields  $|\partial_t W|_{X^{s-1}} \lesssim \frac{1}{\varrho} |W|_{X^s}$ , which is not sufficient to conclude. Thus we need now to use precisely the structure of our system, and in particular the estimate (3.23). Thus we decompose into two components:

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

Let us start with the “slow” contribution,  $[\Lambda^s, T[V]] \Pi \partial_t W$ . One can use equation (3.24) in order to control  $\Pi \partial_t W$ , uniformly with respect to  $\varrho$  small. 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

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

where we used estimate (3.21) in Lemma 3.8, and the property  $\|\Pi L_\varrho\| = \mathcal{O}(\varrho)$ . It follows

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

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

We continue with the “fast” contribution,  $[\Lambda^s, T[V]] (\text{Id} - \Pi) \partial_t W$ . Since  $(\text{Id} - \Pi)$  is constant, it commutes with  $\Lambda^s$ , and Kato-Ponce's commutator estimates (Lemma A.3) yield

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

Now, one has as above,

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

with  $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$ . Estimate (3.23) in Lemma 3.8 allows to recover a factor of size  $\mathcal{O}(\varrho)$ :

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



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

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

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

Altogether, one has, applying Cauchy-Schwarz inequality,

$$|([\Lambda^s, T[V]] \partial_t W, \Lambda^s W)| \leq C_0 M (|W|_{X^s} + |\mathcal{R}|_{X^{s-1}}) |W|_{X^s}, \quad (3.31)$$

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

Plugging (3.27), (3.28), (3.29), (3.30), (3.31) into (3.26) yields

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

Finally, estimate (3.20) 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 Gronwall-Bihari's Lemma.  $\square$

**Completion of the proof of Theorem 1.2.** Let us now quickly show how Theorem 1.2 follows from Lemma 3.9. For a 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 (3.1) yields  $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top \in C([0, T_{\max}); X^{s+1}) \cap C^1([0, T_{\max}); X^s)$  solution to (3.3).

Thanks to Proposition 3.4,  $V_{\text{app}} = V_{\text{RL}} + V_{\text{em}}$  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

$$\sup_{t \in [0, T/M]} \{ |V_{\text{app}}(t, \cdot)|_{X^{s+1}} + |\partial_t V_{\text{app}}(t, \cdot)|_{X^s} \} \leq C_1 M. \quad (3.32)$$

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

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

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

$$T^\sharp \equiv \sup \left\{ t \in [0, T_{\max}, T/M], \|W\|_{L^\infty([0, t]; X^s)} + \varrho \|\partial_t W\|_{L^\infty([0, t]; X^{s-1})} \leq 2C_2 \varrho M \right\}. \quad (3.33)$$

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

Recall that  $W$  satisfies (3.24); thus we apply Lemma 3.9 with

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

Proposition 3.4 yields

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

Now, using that  $(X^s, |\cdot|_{X^s})$  is a Banach algebra, and using (3.19) in Lemma 3.7, one has

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

It follows from the above estimates that

$$\|\mathcal{R}\|_{L^\infty([0, T^\sharp]; X^s)} \leq C_3 (M^2 \varrho + M \varrho^2), \quad (3.34)$$

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

Finally, we apply (3.25) in Lemma 3.9 (making use of (3.32),(3.33),(3.34)), and deduce

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

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 (3.24):

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

It follows that there exists  $T' > 0$ , depending only 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 (3.32),(3.33) immediately yield

$$\|\zeta_2\|_{L^\infty([0, T^\sharp]; H^s)} + \|u_s\|_{L^\infty([0, T^\sharp]; H^s)} \leq M \exp(C_0 M t), \quad (3.35)$$

$$\|\zeta_1\|_{L^\infty([0, T^\sharp]; H^s)} + \|m\|_{L^\infty([0, T^\sharp]; H^s)} \leq M \varrho \exp(C_0 M t), \quad (3.36)$$

$$\| |\partial_t \zeta_1| + |\partial_t \zeta_2| + |\partial_t u_s| + |\partial_t m| \|_{L^\infty([0, T^\sharp]; H^{s-1})} \leq M \exp(C_0 M t), \quad (3.37)$$

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

It follows in particular from (3.37) that for any  $t \in [0, T^\sharp]$ , one has

$$|h_2(t, \cdot) - h_2(0, \cdot)|_{H^{s-1}} \leq \left| \int_0^t \partial_t \zeta_2(t', \cdot) dt' \right|_{H^{s-1}} \leq C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1}) M t,$$

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 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 (3.33), (3.35) and (3.36) (using Lemma A.1 and Corollary A.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 which fail to meet the smallness assumption in (1.3)?*

In both cases, as we shall see, the answer will be given through a decomposition between fast and slow modes. Such decomposition is exact in the linear case ( $\epsilon = 0$  in (1.1)) as the 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  $\varrho \rightarrow 0$ , two of these waves (corresponding to the solution of the rigid-lid system, and mainly supported on variables  $\zeta_2, u_s$ ) are traveling with velocity  $c_\pm^s \sim \pm 1$ , while the two other ones (mainly supported on  $\zeta_1, m$ ) are traveling with velocity  $c_\pm^f \sim \pm \sqrt{1 + \delta^{-1}/\varrho}$ .

This decomposition is far from being new. In the literature, the two modes are also often referred as surface/interface modes, or barotropic/baroclinic modes, since the fast mode component share the properties of water-waves for one layer of a fluid of constant mass density [15]. The decomposition is exact in the linear setting, and has been showed to hold approximately in the weakly nonlinear setting; see [13], and references therein. In that case, the smallness of  $\epsilon$  allows to control the coupling

terms between each of the waves (even when additional —small— dispersion terms are included), provided the initial data is 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 already mentioned, such result will rely on a condition of spatial localization of the initial data, that we express through weighted Sobolev spaces.

In Section 4.1, we construct slow and fast mode correctors which allow 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 precise the results stated in Proposition 3.4 and Theorem 1.2 with Proposition 4.1 and Theorem 4.5, respectively. In Section 4.2, we extend the consistency result obtained in Proposition 3.4 to ill-prepared initial data, that is data allowing non-small horizontal momentum and deformation of the surface, and thus involving a leading order slow mode. Unfortunately, we cannot carry on the study of Section 3.3, and deduce the stronger result corresponding to Theorem 1.2 (although numerical simulations are in full agreement with such result). Finally, subsection 4.3 contains numerical simulations illustrating the aforementioned results, and an accompanying discussion.

## 4.1 Improved approximate solution

In this section, we show that one can construct the first-order corrector to the rigid-lid approximate solution displayed in Theorem 1.2, provided the initial data is bounded in weighted Sobolev spaces. A key ingredient is the establishment of a fast mode corrector, which allows to take into account small initial data supported on variables  $\zeta_1, m$ .

In Proposition 4.1, 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 Section 3, and one obtains the stronger result expressed in Theorem 4.5, below.

**Proposition 4.1.** *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 (1.3),(1.4) (after the change of variable (3.2)) with given  $0 < M, h_0 < \infty$ . Assume additionally that there exists  $\sigma > 1/2$  such that*

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

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

$$\forall t \in [0, T/M], \quad |V_{\text{RL}}|_{X^{s+1}} + |\partial_t V_{\text{RL}}|_{X^s} \leq C_0 M .$$

2.  $V_{\text{cor}}^s \equiv (\varrho \check{\zeta}_1, 0, 0, 0)^\top$  is well-defined with

$$\check{\zeta}_1 \equiv -\left(\eta + \frac{\delta}{2}\eta^2\right) - \frac{(1-\eta)(\delta^{-1} + \eta)v^2}{\gamma(1 + \delta^{-1})^2} .$$

3.  $V_{\text{cor}}^f$  is well-defined with

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

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

4. There exists  $V_{\text{rem}}$ , with

$$\forall t \in [0, T], \quad |V_{\text{rem}}(t, \cdot)|_{X_{\text{ul}}^{s+1}} \leq C_0 M ,$$

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

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

**Remark 4.2.** *The last item is restricted to the time domain  $t \in [0, T]$  instead of  $t \in [0, T/M]$ . As discussed in Remark 1.3, we do not expect any particular limitation when  $M$  is small.*

**Remark 4.3.** *We denote  $(H_{\text{ul}}^s, |\cdot|_{H_{\text{ul}}^s})$  the uniformly local Sobolev space introduced in [20]:*

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

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

We then denote  $(X_{\text{ul}}^s, |\cdot|_{X_{\text{ul}}^s})$  and  $(L^\infty([0, T]; X_{\text{ul}}^s), \|\cdot\|_{L^\infty([0, T]; X_{\text{ul}}^s)})$  similarly to the previously defined Sobolev-based spaces.

*Proof of Proposition 4.1.* The well-posedness and estimate of  $V_{\text{RL}}$  for  $t \in [0, T/M]$  has already been stated in Proposition 3.4. 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  $\zeta_2^0, u_s^0$ , using the corresponding solution to the rigid-lid system, and that we will refer as *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 regards 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 the 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 reach the improved precision. More precisely, we seek  $V_{\text{app}}^s \equiv V_{\text{RL}} + V_{\text{cor}}^s + \varrho^2 V_{\text{rem}}$ , with  $V_{\text{RL}} + V_{\text{cor}}^s \equiv (\varrho \check{\zeta}_1, \eta, v, 0)^\top$  as in the proof of Proposition 3.4, and  $V_{\text{rem}} \equiv (0, 0, 0, \check{m})^\top$  to be determined. Following the exact same steps as in the proof of Proposition 3.4, we see that the only difficulty we face lies in the estimate of

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

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

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

where we denote  $\check{h}_1 \equiv 1 - \eta$  and  $h_2 \equiv \delta^{-1} + \eta$ .

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

$$\forall t \in [0, T/M], \quad |V_{\text{RL}}|_{X^{s+1}} + |\partial_t V_{\text{RL}}|_{X^s} \lesssim C_0 M, \quad (4.2)$$

$$\forall t \in [0, T/M], \quad |\check{\zeta}_1|_{H^{s+1}} + |\partial_t \check{\zeta}_1|_{H^s} \lesssim C_0 M. \quad (4.3)$$

(here and below, we denote  $C_0 = C(M, h_0^{-1}, \delta_{\text{min}}^{-1}, \delta_{\text{max}}, \gamma_{\text{min}}^{-1})$ ). One deduces

$$\forall t \in [0, T/M], \quad |\check{m}|_{H_{\text{ul}}^{s+1}} + |\partial_x \check{m}|_{H^s} \lesssim C_0 M, \quad (4.4)$$

where we use that  $H^s$  is continuously embedded in  $H_{\text{ul}}^s$  and  $H_{\text{ul}}^s$  is a Banach algebra, for any  $s \geq s_0$  (see, e.g., [25, App. B.4]). The estimate on  $V_{\text{rem}}$ , stated in the Proposition, is given by (4.2),(4.3),(4.4).

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

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

where we used the commutator estimate recalled in Lemma A.3. Using the above, it is now straightforward to check that  $V_{\text{app}}^s \equiv V_{\text{RL}} + V_{\text{cor}}^s + \varrho^2 V_{\text{rem}} \equiv (\varrho \check{\zeta}_1, \eta, v, \varrho^2 \check{m})^\top$  satisfies (3.3), up to a remainder term,  $R^s$ , with

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

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

*Construction of the fast mode approximate solution.* The corrector  $V_{\text{cor}}^f$  has been defined as the unique solution to

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

with initial data  $V_{\text{cor}}^f|_{t=0} \equiv (\zeta_1^0 - \varrho \check{\zeta}_1|_{t=0}, 0, 0, m^0)^\top$ .

Our aim is to prove that  $V_{\text{cor}}^f$  is an approximate solution to (3.3). We recall that the system reads

$$\partial_t V + \frac{1}{\varrho} (L_\varrho + \varrho B[V]) \partial_x V = 0, \quad \text{with } L_\varrho \equiv \begin{pmatrix} 0 & 0 & \frac{\gamma-1}{\gamma(\delta+1)} & \frac{\gamma+\delta}{\gamma(\delta+1)} \\ 0 & 0 & \frac{\varrho}{1+\delta} & \frac{\varrho}{1+\delta} \\ 0 & \varrho(\gamma+\delta) & 0 & 0 \\ \gamma(1+\delta^{-1}) & \varrho \frac{\delta+\gamma}{\delta} & 0 & 0 \end{pmatrix}.$$

Thus  $V_{\text{cor}}^f$  satisfies

$$\partial_t V_{\text{cor}}^f + \frac{1}{\varrho} (L_\varrho + \varrho B[V_{\text{cor}}^f]) \partial_x V_{\text{cor}}^f = R^f,$$

with

$$R^f \equiv \frac{1}{\varrho} (L_\varrho - L_{(0)}) \partial_x V_{\text{cor}}^f + B[V_{\text{cor}}^f] \partial_x V_{\text{cor}}^f.$$

It is obvious that for any  $t \in \mathbb{R}$ ,  $V_{\text{cor}}^f$  satisfies

$$|V_{\text{cor}}^f(t, \cdot)|_{X^{s+1}} \lesssim |V_{\text{cor}}^f|_{t=0}|_{X^{s+1}} \leq C_0 M \varrho, \quad (4.7)$$

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

In particular, Lemma 3.7 and Lemma A.1 yield

$$|B[V_{\text{cor}}^f] \partial_x V_{\text{cor}}^f|_{X^s} \lesssim |V_{\text{cor}}^f|_{L^\infty(\mathbb{R}^4)} |V_{\text{cor}}^f|_{X^{s+1}} \leq C_0 M^2 \varrho^2. \quad (4.8)$$

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

It is straightforward to check that

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

so that

$$\left| \frac{1}{\varrho} (L_\varrho - L_{(0)}) \partial_x V_{\text{cor}}^f \right|_{X^s} = \left| \frac{1}{\varrho} (L_\varrho - L_{(0)})(\text{Id} - \Pi) \partial_x V_{\text{cor}}^f \right|_{X^s} \lesssim C_0 M \varrho^2. \quad (4.9)$$

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

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

*Completion of the proof.* One easily checks that  $V_{\text{app}} \equiv V_{\text{app}}^s + V_{\text{cor}}^f$  satisfies

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

where

$$R^c \equiv (B[V_{\text{app}}] - B[V_{\text{cor}}^f])\partial_x V_{\text{cor}}^f + (B[V_{\text{app}}] - B[V_{\text{app}}^s])\partial_x V_{\text{app}}^s.$$

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

Note that similarly to (3.19) in Lemma 3.7, one can check that estimates (4.2) (4.3), (4.4), (4.5) and (4.7) yield

$$|R^c|_{X^s} \leq C_0 \times (\|V_{\text{RL}} \otimes \partial_x V_{\text{cor}}^f\|_{X^s} + \|V_{\text{cor}}^f \otimes \partial_x V_{\text{RL}}\|_{X^s} + M \varrho^2) \quad (4.11)$$

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  $V$ .

In order to control the latter contribution, we make use of the fact that the initial data is assumed to be spatially localized. Thus  $V_{\text{cor}}^f$  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  $V_{\text{RL}}$  remains spatially localized around  $x = 0$  on the relevant timescale.

It is convenient here to restrict ourselves to the time domain to  $t \in [0, T]$  instead of the more stringent  $t \in [0, T/M]$ . Indeed, recall the linearized equation predicts that  $V_{\text{RL}}$  is the superposition of two counter-propagating waves with velocity  $\pm 1$ , so that the spatial localization does not hold for large time (small  $M$ ). One could work with a time-varying spatial localization (*e.g.* replacing  $\langle x \rangle^\sigma \equiv (1 + |x|^2)^{\sigma/2}$  with  $(1 + |x/\langle t \rangle|)^\sigma$ ) but again, as our analysis is focused on strong nonlinearities, it is easier to assume  $M \approx 1$  and restrict ourselves to  $t \in [0, T]$ .

We state and prove below the persistence of the spatial decay which holds generically for a quasilinear, hyperbolic system; and complete the proof of Proposition 4.1 thereafter.

**Lemma 4.4** (Persistence of spatial decay). *Let  $s \geq s_0 + 1$ ,  $s_0 > 1/2$  and  $V_{\text{RL}} \equiv (\eta, v)^\top$  be the solution to (1.2), with initial data  $V_{\text{RL}}|_{t=0} \equiv (\eta^0, v^0)^\top$  as above. 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 denote  $\langle x \rangle \equiv (1 + |x|^2)^{1/2}$ ). There exists  $M > 0$  such that if  $|(\eta^0, v^0)^\top|_{H^s \times H^s} \leq M$ , then one has*

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

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

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

where  $S[\cdot], \Sigma[\cdot]$  are smooth mappings onto the space of 2-by-2 symmetric matrices ( $S$  and  $\Sigma$  are explicit; see [17] for more details).

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

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

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

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

Using the usual technique for *a priori*  $H^s$  estimates (see Lemma A.7 for example), one obtains

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

Now, using the control of  $V_{\text{RL}} \in X^s$  in (4.2), and since one has

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

it follows from Gronwall-Bihari's inequality:

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

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

Let us now complete the proof of Proposition 4.1. We use the following calculation to estimate  $R^c$  in (4.11). Set  $s > 1/2$ ,  $\sigma \geq 0$ , 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 one has

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

and one can check (see [23] for example) that for any  $\sigma > 1/2$  and  $T > 0$ , one has

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

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

It is now straightforward, applying Lemma 4.4, the definition of  $V_{\text{cor}}^f$ , (4.7) and the above calculations to (4.11), that the following estimate holds:

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

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) conclude the proof of Proposition 4.1.  $\square$

Let us conclude this section with the following result, which corresponds to Theorem 1.2, when Proposition 4.1 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 (1.4) holds with  $h_0 > 0$  and there exists  $0 < M < \infty$  and  $\sigma > 1/2$  such that*

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

and

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

Then there exists  $T^{-1}, C$ , depending uniquely 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  $V_{\text{RL}}, V_{\text{cor}}^s, V_{\text{cor}}^f$  as in Proposition 4.1. Denote  $U_{\text{app}}$  the approximate solution corresponding to  $V_{\text{RL}} + V_{\text{cor}}^s + V_{\text{cor}}^f$ , after the change of variables (3.2). Then one has

$$\|U - U_{\text{app}}\|_{L^\infty([0, T]; X_{\text{ui}}^s)} \leq C M \varrho^2.$$

**Remark 4.6.** *Recall we set  $\epsilon = 1$  and  $\alpha = \varrho$  after Theorem 1.2; see Remarks 1.3 and 1.4. The statement of Theorem 4.5 is easily extended to the general case.*

*Sketch of the proof.* The existence and uniqueness of  $U$  has already been stated in Theorem 1.2. The existence and uniqueness of  $V_{\text{RL}}, V_{\text{cor}}^s, V_{\text{cor}}^f$  is guaranteed by Proposition 4.1. Now, one can follow the exact same procedure as described in Section 3 (and especially Section 3.3), using the result of Proposition 4.1 instead of the corresponding Proposition 3.4. Note however that the remainder term constructed in Proposition 4.1,  $V_{\text{rem}}$ , may not have 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 [20], the usual theory of hyperbolic quasilinear system in Sobolev spaces extends naturally to uniformly local Sobolev spaces, without significant change in the proof (in particular, similar product and commutator estimates hold; see [25, App. B]); thus we do not detail further on.

We simply remark that  $V_{\text{app}}$  has been constructed so that  $W \equiv V - V_{\text{app}}$  satisfies

$$|W|_{t=0}|_{X_{\text{ui}}^s} \lesssim C_0 M \varrho^2,$$

where we denote  $V \equiv (\zeta_1, \zeta_2, u_s, m)^\top$  the solution to (3.3) corresponding to  $U$ , in terms of the variables defined by (3.1). Consequently, the energy estimate (3.25) in Lemma 3.9 implies

$$\forall t \in [0, T], \quad |W|_{X_{\text{ui}}^s} \lesssim C_0 M \varrho^2 + \int_0^t |R(t', \cdot)|_{X_{\text{ui}}^s} dt',$$

and Proposition 4.1 immediately yields the desired estimate.  $\square$

## 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 which fail to meet the smallness assumption in (1.3), or in other words *admitting a non-small 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, that we shall exhibit below. There are two main differences with 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 non-trivial contribution on the behavior of the fast mode approximate solution, and cannot be neglected.
2. The strategy developed in Section 3 is not valid anymore, as the hypothesis of Lemma 3.9 is no longer satisfied.

As a consequence, we restrict our statement to a consistency result, namely Proposition 4.7, below; we cannot deduce an estimate on the difference between the exact and the approximate solution, as in Theorems 1.2 and 4.5, or even prove that (1.1) is well-posed on a time interval independent of  $\varrho$  small. However, numerical simulations, presented in the subsequent subsection, are in full agreement with the intuitive conjecture that

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

with the notations introduced below.

**Proposition 4.7.** *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 (1.4) (after the change of variable (3.2)) with given  $h_0 > 0$ . Assume additionally that there exists  $0 < M < \infty$  and  $\sigma > 1/2$  such that*

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

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

$$\forall t \in [0, T/M], \quad |V_{\text{RL}}|_{X^{s+2}} + |\partial_t V_{\text{RL}}|_{X^{s+1}} \leq C_0 M.$$

2.  $V_{\text{cor}}^f$  is well-defined with

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

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

3. There exists  $V_{\text{rem}}$  with

$$\forall t \in [0, T], \quad |V_{\text{rem}}|_{X^{s+1}} + \varrho |\partial_t V_{\text{rem}}|_{X^s} \leq C_0 M,$$

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

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



**Remark 4.8.** *The fast mode contribution  $V_{\text{cor}}^f$  is different from the one defined in Proposition 4.1. Moreover, it is not a corrector term per se, since it has the same order of magnitude as  $V_{\text{RL}}$ . We decide to use the same notation in order to acknowledge the following fact: one can replace  $V_{\text{cor}}^f$  in Proposition 4.1 by the one defined above, 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 4.7.* We follow the same three steps as in the proof of Proposition 4.1. We first construct an approximate solution corresponding to the slow mode and fast mode, respectively. Finally, we prove that the coupling effects between the two modes are weak, thanks to the appropriate spatial localization of the initial data, and therefore the superposition of the two modes yields an approximate solution.

*Construction of the slow mode approximate solution.* Proposition 3.4 directly gives the desired result: denoting  $V_{\text{rem}}^s \equiv (\check{\zeta}_1, 0, 0, \varrho \check{m})$ , with  $\check{\zeta}_1, \check{m}$  defined in (3.14), (3.16), one has

$$\forall t \in [0, T/M], \quad |V_{\text{RL}}|_{X^{s+2}} + |\partial_t V_{\text{RL}}|_{X^{s+1}} \lesssim C_0 M, \quad (4.15)$$

$$\forall t \in [0, T/M], \quad |V_{\text{rem}}^s|_{X^{s+2}} + |\partial_t V_{\text{rem}}^s|_{X^{s+1}} \lesssim C_0 M, \quad (4.16)$$

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

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

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

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

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

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

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

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

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

In the following, we seek an approximate solution to

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

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

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

$$V_{\text{app}}^f(t, x) = V_{\text{cor}}^f(t, t/\varrho, x) + \varrho V_{\text{rem}}^f(t, t/\varrho, x),$$

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

$$\frac{1}{\varrho} \partial_\tau V_{\text{app}}^f + \partial_t V_{\text{app}}^f + \left( \frac{1}{\varrho} L_{(0)} + L_{(1)} + B_{(1)}[(\text{Id} - \Pi)V_{\text{app}}^f] \right) \partial_x V_{\text{app}}^f = 0. \quad (4.19)$$

Based on the fact that at first order (in terms of  $\varrho$ ), 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  $V_{\text{cor}}^f$  as the superposition of decoupled waves, supported on the eigenvectors of  $L_{(0)}$  corresponding to non-zero eigenvalues.

The analysis of higher-order terms yields

- the behavior of  $V_{\text{cor}}^f$  with respect to the large-time variable,  $t$ , which takes into account the nonlinear effects on the propagation of each decoupled waves;
- a remainder term,  $V_{\text{rem}}^f(t, \tau, x)$ , which mimics the coupling effects between the two counter-propagating waves of  $V_{\text{cor}}^f$ , as well as the “slow mode component”,  $\Pi V_{\text{cor}}^f$ .

The key ingredient in the proof is to show that one can set  $V_{\text{cor}}^f$  such that  $V_{\text{rem}}^f$  remains small for large time. 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 [5, 29], and later on to similar problems in the bi-fluidic setting [13, 14]. The strategy is described comprehensively for example in [25, Chap. 7], thus we do not detail the calculations, and simply state the outcome.

It is convenient to introduce here the following eigenvectors of  $L_{(0)}$ :<sup>3</sup>

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

We set

$$V_{\text{cor}}^f(\cdot, \tau, x) \equiv u_+(\cdot, x - c\tau)\mathbf{e}_+ + u_-(\cdot, x + c\tau)\mathbf{e}_-,$$

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  $u_{\pm}|_{t=0} = \frac{1}{2}(\zeta_1^0 \pm c^{-1}m^0)$ . One checks immediately that  $V_{\text{cor}}^f : (t, x) \mapsto V_{\text{cor}}^f(t, t/\varrho, x)$  is as in the Proposition, explaining our (slightly misused) notation.

In the same way, we write

$$V_{\text{rem}}^f(\cdot, \tau, x) \equiv r_+(t, \tau, x)\mathbf{e}_+ + r_-(t, \tau, x)\mathbf{e}_- + r_0(t, \tau, x)\mathbf{e}_0,$$

with functions  $r_+, r_-, r_0$  determined by

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

and  $V_{\text{rem}}^f(\cdot, 0, x) \equiv 0$ .

One can check that  $V_{\text{app}}^f(t, \tau, x) = V_{\text{cor}}^f(t, \tau, x) + \varrho V_{\text{rem}}^f(t, \tau, x)$ , as defined above, satisfies

$$\frac{1}{\varrho} \partial_{\tau} V_{\text{app}}^f + \partial_t V_{\text{app}}^f + \left( \frac{1}{\varrho} L_{(0)} + L_{(1)} + B_{(1)}[V_{\text{app}}^f] \right) \partial_x V_{\text{app}}^f = R^f,$$

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

It follows (using (3.19) in Lemma 3.7) that

$$|R^f|_{X^s} \leq \varrho C(|\partial_t V_{\text{rem}}^f|_{X^s}, |V_{\text{rem}}^f|_{X^{s+1}}, |V_{\text{cor}}^f|_{X^{s+1}}). \quad (4.20)$$

In order to estimate the above, one needs to control  $V_{\text{rem}}^f$ , using the following two Lemmata.

<sup>3</sup>Of course a fourth vector—second linearly independent element of  $\ker(L_{(0)})$ —could be defined, but this is not necessary in our analysis.

**Lemma 4.9.** *Let  $s \geq 0$ , and  $f^0 \in H^s(\mathbb{R})$ . Then there exists a unique global strong solution,  $u(\tau, x) \in C^0(\mathbb{R}; H^{s+1}) \cap C^1(\mathbb{R}; H^s)$ , of*

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

where  $c_1 \neq c_2$ . Moreover, one has the following estimates for any  $\tau \in \mathbb{R}$ :

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

**Lemma 4.10.** *Let  $s \geq s_0 > 1/2$ , 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*

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

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

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

Lemma 4.9 is straightforward, and Lemma 4.10 follows from Proposition 3.5 in [23].

Lemmata 4.9 and 4.10 applied to  $V_{\text{rem}}^f$  immediately yield

$$|V_{\text{rem}}^f(t, \tau, \cdot)|_{X^{s+1}} \leq C |u_\pm(t, \cdot)|_{H^{s+1}(\mathbb{R})} + C |u_+(t, \cdot)(1 + |\cdot|^2)^\sigma|_{H^{s+2}(\mathbb{R})} |u_-(t, \cdot)(1 + |\cdot|^2)^\sigma|_{H^{s+2}(\mathbb{R})}.$$

One can apply the same arguments to  $\partial_t V_{\text{rem}}^f$  (deriving 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 one obtains

$$|\partial_t V_{\text{rem}}^f(t, \tau, \cdot)|_{X^s} \leq C |u_\pm(t, \cdot)|_{H^{s+1}(\mathbb{R})} + C |u_+(t, \cdot)(1 + |\cdot|^2)^\sigma|_{H^{s+2}(\mathbb{R})} |u_-(t, \cdot)(1 + |\cdot|^2)^\sigma|_{H^{s+2}(\mathbb{R})}.$$

It is not difficult to show that the inviscid Burgers' equation propagates the localization in space of its solutions (see Lemma 4.4), so that one has

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

thus we proved

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

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

Finally, we recall that  $V_{\text{cor}}^f \equiv u_+(t, x - ct/\varrho) \mathbf{e}_+ + u_-(t, x + ct/\varrho) \mathbf{e}_-$  and  $V_{\text{rem}}^f \equiv V_{\text{rem}}^f(t, t/\varrho, x)$ , and one deduces

$$\forall t \in [0, T], \quad |V_{\text{cor}}^f|_{X^{s+2}} + \varrho |\partial_t V_{\text{cor}}^f|_{X^{s+1}} \leq C_0 M, \quad (4.22)$$

$$\forall t \in [0, T], \quad |V_{\text{rem}}^f|_{X^{s+1}} + \varrho |\partial_t V_{\text{rem}}^f|_{X^s} \leq C_0 M, \quad (4.23)$$

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

Therefore (4.20) simply becomes

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

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

*Completion of the proof.* One easily checks that  $V_{\text{app}} \equiv V_{\text{app}}^s + V_{\text{app}}^f \equiv V_{\text{RL}} + V_{\text{cor}}^f + \varrho V_{\text{rem}}^s + \varrho V_{\text{rem}}^f$  satisfies

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

where  $R^s$  and  $R^f$  have been defined and estimated above, and with

$$R^c \equiv (B[V_{\text{app}}] - B[V_{\text{app}}^f]) \partial_x V_{\text{app}}^f + (B[V_{\text{app}}] - B[V_{\text{app}}^s]) \partial_x V_{\text{app}}^s.$$

The contribution of  $R^f + R^s$  is controlled as a result of the above calculations; see (4.17) and (4.24). Thus the only component to control comes from the coupling effects between  $V_{\text{app}}^s$  and  $V_{\text{app}}^f$ , displayed in  $R^c$ . Recalling the construction of  $V_{\text{app}}^s \equiv V_{\text{RL}} + \varrho V_{\text{rem}}^s$  and  $V_{\text{app}}^f \equiv V_{\text{cor}}^f + \varrho V_{\text{rem}}^f$ , using estimates (4.15), (4.16), (4.22), (4.23), one can check that

$$\|R^c\|_{X^s} \leq C_0 \times (\|V_{\text{RL}} \otimes \partial_x V_{\text{cor}}^f\|_{X^s} + \|V_{\text{cor}}^f \otimes \partial_x V_{\text{RL}}\|_{X^s} + M\varrho),$$

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 above as in the proof of Proposition 4.1, using spatial localization. For any function  $v$  satisfying  $(1 + |\cdot|^2)^\sigma v(t, \cdot) \in H^s$ , one has

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

and recall that for any  $\sigma > 1/2$  and  $t > 0$ , one has

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

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

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

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

with  $C_0 = C(M, h_0^{-1}, \frac{1}{2\sigma-1}, \delta_{\min}^{-1}, \delta_{\max}, \gamma_{\min}^{-1})$ . Proposition 4.7 is proved.  $\square$

**Remark 4.11.** *As mentioned previously, we are unable to deduce from Proposition 4.7 a rigorous estimate on the difference between the exact solution and the constructed approximate solution as in Theorems 1.2 or 4.5. Indeed, the strategy developed in Section 3.3 fails, as the solution does not satisfy the assumption of Lemma 3.9, and 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  $\|[\partial_t, T[V]] \Pi^s W\|_{L^2}$ ; and even more precisely  $\|[\partial_t, T[V]] \Pi \Lambda^s W\|_{L^2}$ , as the complement is estimated through (3.23) in Lemma 3.8. We expect that the following strategy would imply the desired result: decompose*

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

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

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

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

### 4.3 Discussion and numerical simulations

In this section, we illustrate and discuss the results displayed in Theorem 1.2 and Proposition 3.4 (validity of the rigid-lid approximation), Proposition 4.1 and Theorem 4.5 (improved approximate solution) and Proposition 4.7 (case of ill-prepared initial data).

In each case, we construct the appropriate approximate solution  $(V_{\text{RL}}, V_{\text{cor}}^f, V_{\text{cor}}^s)$  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  $\varrho$  (and  $\alpha = \varrho$ ), while the other parameters are fixed.

More precisely, we set:

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

The initial data is set as follows:

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

and

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

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  $1 + \delta^{-1} + \epsilon \zeta_1|_{t=0}$ ,  $\delta^{-1} + \epsilon \zeta_2|_{t=0}$ ,  $1 + u_s|_{t=0}$  and  $m|_{t=0}$ . The lower panel represents the solution of the free-surface system (3.3) as well as the corresponding approximate solution at stake (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 log-log scale, for several values of  $\varrho$  (the markers reveal the positions which have been computed), at final time  $T = 4$ .

The numerical scheme we use is based on spectral methods as for the space discretization (see [33]), thus yields an exponential accuracy with respect to the size of the grid  $\Delta x$ , 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 [30], with a tolerance of  $10^{-8}$ , in order to solve the time-dependent problem.

**Well-prepared initial data.** In Figure 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). One straightforwardly sees 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  $\varrho$  (we recall  $\gamma = 0.1$  in the panel 2(c)). As a matter of fact, the precision of the approximation is not foreseen from Theorem 1.2: as we can see from the panel 2(b), the convergence rate for  $\zeta_2$  and  $u_s$  is  $\mathcal{O}(\varrho^2)$ , as Theorem 1.2 predicts only  $\mathcal{O}(\varrho)$ . One can see that the main error in the rigid-lid approximation is supported on the deformation of the surface,  $\zeta_1$ , as well as on the horizontal momentum,  $m$  (and more precisely the fast mode of the horizontal momentum).

Of course, such result is predicted by Theorem 4.5, since the first-order corrector constructed in Proposition 4.1 follows precisely the above description. We show in Figure 3 the precision of the improved rigid-lid approximation. One sees 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  $\zeta_1, \zeta_2, u_s, m$ , in full accordance with Theorem 4.5.

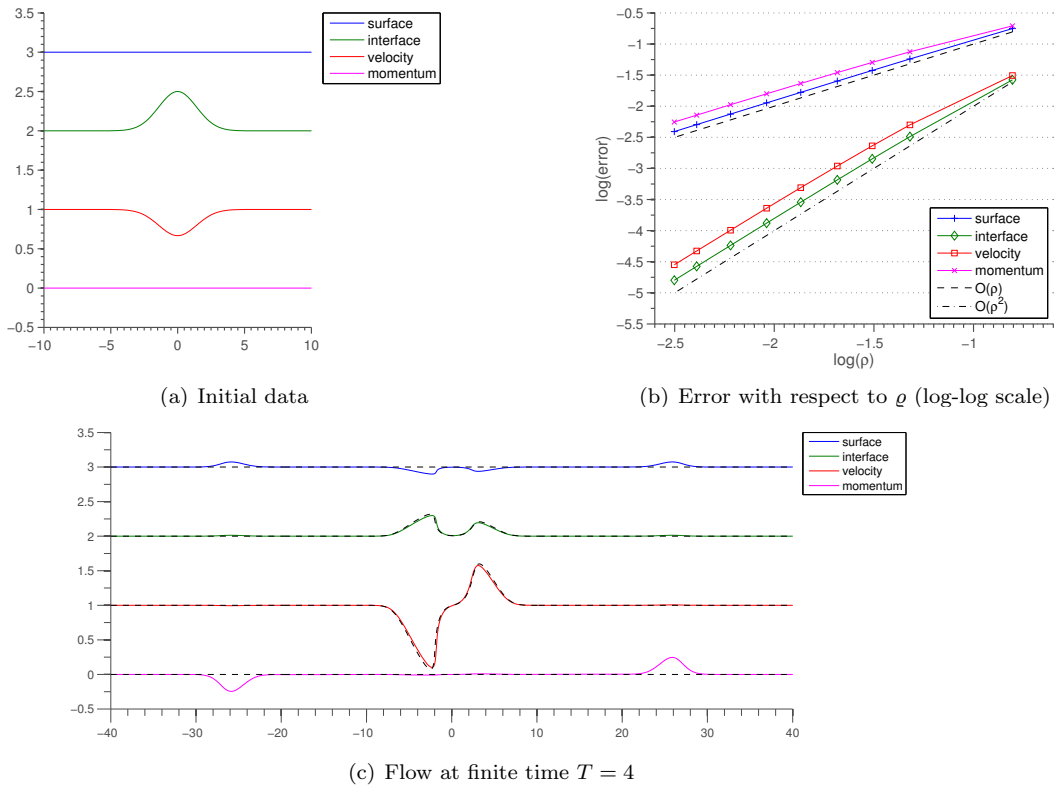


Figure 2: Solution of the free-surface system compared with the rigid-lid approximate solution

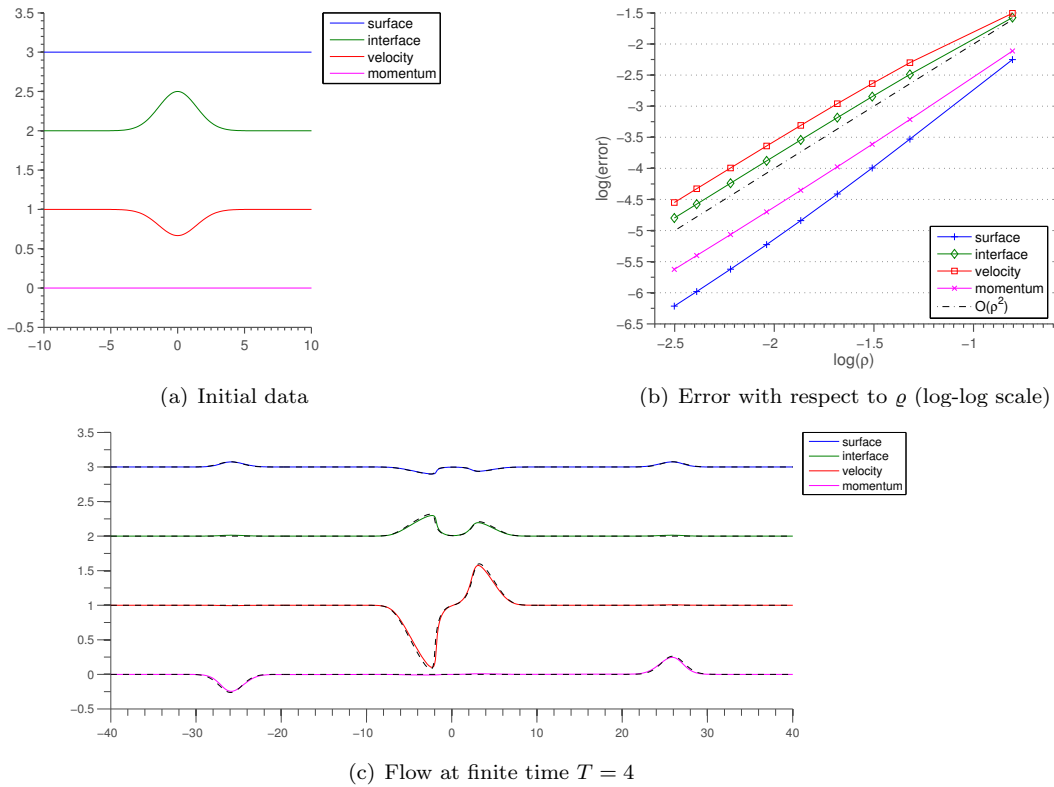


Figure 3: Solution of the free-surface system compared with the improved approximate solution

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

We plot in Figure 4 the difference between the exact solution of the free-surface system and the approximate solution constructed in Proposition 4.7. As one can see, 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  $\zeta_1, m$ . As a matter of fact, the variables  $\zeta_2, u_s$  present a slightly better convergence rate in panel 4(b) (around  $\mathcal{O}(\varrho^{1.2})$  and  $\mathcal{O}(\varrho^{1.5})$ , respectively) than predicted by Proposition 4.7, namely  $\mathcal{O}(\varrho)$ .

Such a result advocates for the construction of a higher-order approximation, similarly to the case of well-prepared initial data. Indeed, we know from Proposition 4.1 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 has to modify the initial data of the fast mode corrector in order to ensure that the full approximate solution enjoys the appropriate initial data. Using both statements of Proposition 4.1 and Proposition 4.7, we define the improved approximation for ill-prepared initial data as

$$V_{\text{app}} = V_{\text{RL}} + V_{\text{cor}}^s + V_{\text{cor}}^f,$$

where

- $V_{\text{RL}} \equiv (0, \eta, v, 0)^\top$  is defined by Definition 3.3;
- $V_{\text{cor}}^s \equiv (\varrho\check{\zeta}_1, 0, 0, 0)^\top$  is defined by  $\check{\zeta}_1 \equiv -(\eta + \frac{\delta}{2}\eta^2) \frac{(1-\eta)(\delta^{-1}+\eta)v^2}{\gamma(1+\delta^{-1})^2}$ .
- $V_{\text{cor}}^f$  is defined with

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

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  $u_\pm|_{t=0} = \frac{1}{2}(\zeta_1^0 - \varrho\check{\zeta}_1|_{t=0} \pm c^{-1}m^0)$ .

Let us notice that, as previously mentioned in Remark 4.8, this improved approximation is equivalent to the one already defined in Proposition 4.1 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 4.1) as well as in the ill-prepared case (Proposition 4.7).

We investigate in Figure 5 the accuracy of this improved approximate solution. Comparing panels 4(c) and 5(c), one clearly sees that the new approximate solution shows a better resemblance than the original approximate solution; the main discrepancy seems to be recovered. However, as one can see from panel 5(b), this apparent improvement is not reflected in the convergence rate. Although the produced error is clearly smaller, the rate is not better than  $\mathcal{O}(\varrho)$  where  $\zeta_1$  and  $m$  are involved ( $\zeta_2$  and  $u_s$  are unchanged). It is not clear whether a better approximate solution can be constructed, nor what explains the slightly better convergence rate on  $\zeta_2$  and  $u_s$ . Our numerical simulations indicate that there is a non-trivial coupling between the fast and slow modes during early times (when both are localized at the same place), and that the contribution of these coupling effects is of size  $\approx \varrho$ . Thus in order to take into account these coupling effects, one may have no other choice than solving a fully coupled system, at least for small time,  $t = \mathcal{O}(\varrho)$ .

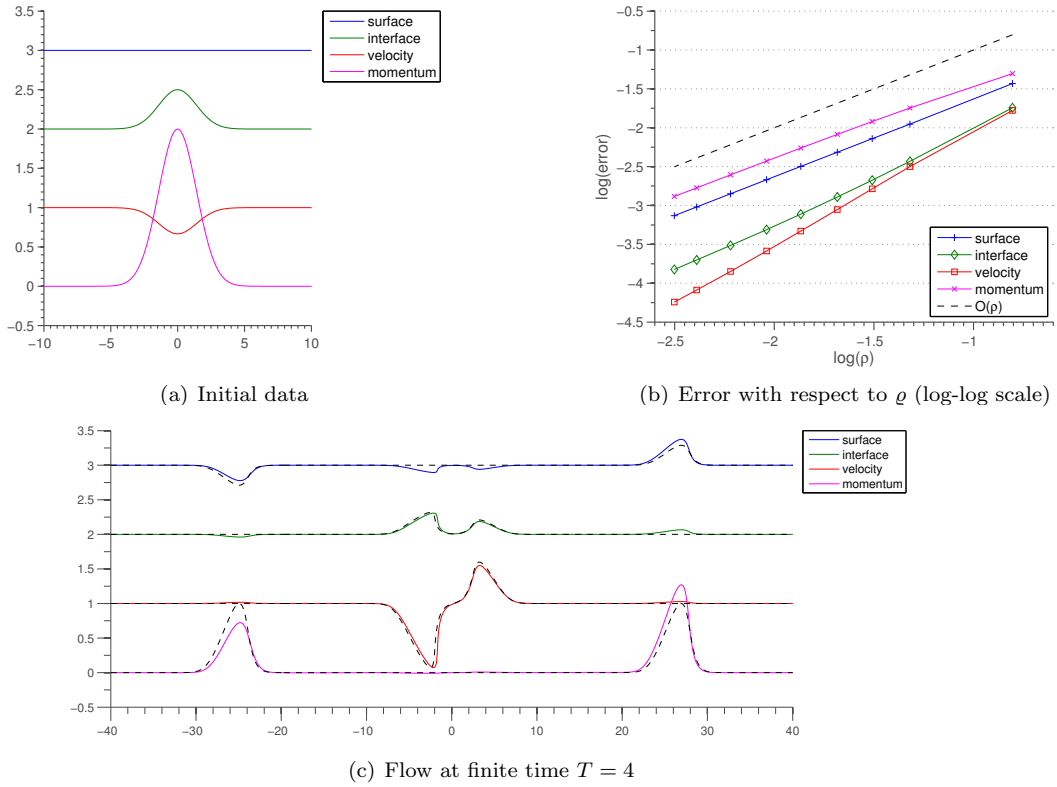


Figure 4: Solution of the free-surface system compared with the approximate solution, for ill-prepared initial data

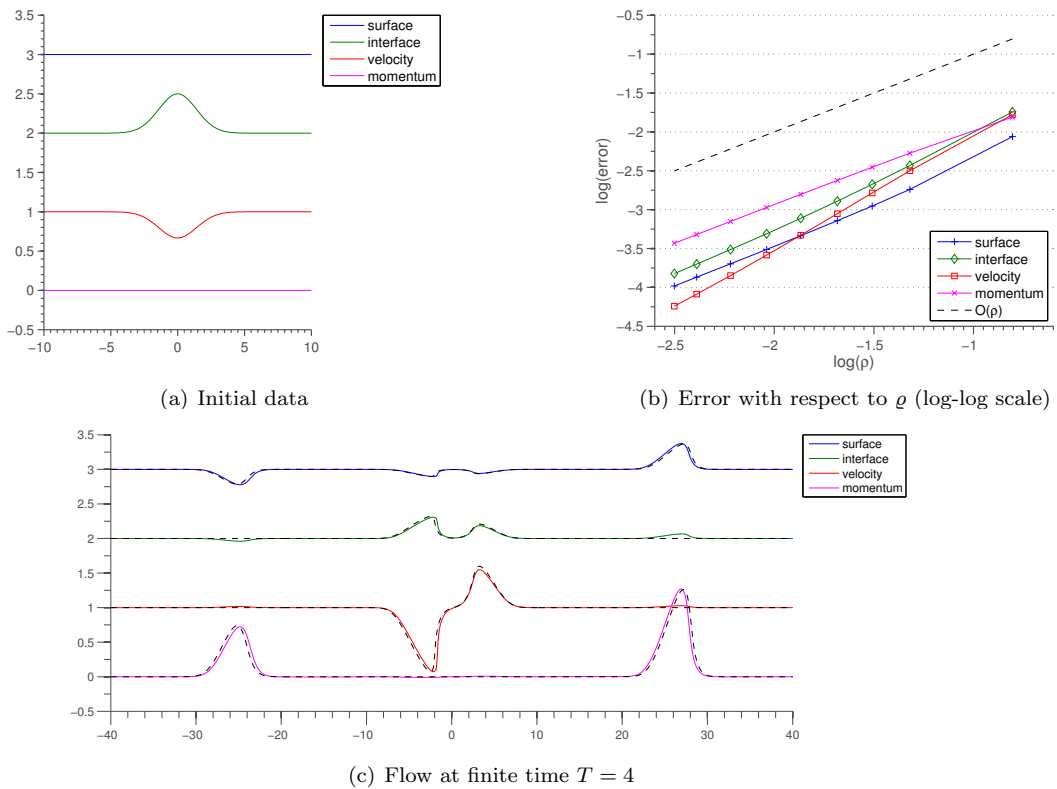


Figure 5: Solution of the free-surface system compared with the improved approximate solution, for ill-prepared initial data



## A 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 let the reader refer to, *e.g.*, [2, 25] for the proof of Lemmata A.1 and A.3.

**Lemma A.1** (Product estimates).

Let  $s \geq 0$ . For any  $f, g \in H^s(\mathbb{R}) \cap L^\infty(\mathbb{R})$ , one has:

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

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

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

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

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

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

**Corollary A.2.** Let  $f, \zeta \in L^\infty \cap H^s$ , 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

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

*Proof.* We will use the identity

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

By Lemma A.1, one deduces

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

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

$$F(X) = \begin{cases} \frac{X}{1-X} & \text{if } 1 - X \geq h_0 > 0, \\ 0 & \text{if } 1 - X \leq 0. \end{cases}$$

The function  $F$  satisfies the hypotheses of Lemma A.1, and one has

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

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.  $\square$

The following Lemma presents a generalization of the Kato-Ponce [22] commutator estimates due to Lannes [24] (one has  $|f|_{H^s}$  instead of  $|\partial_x f|_{H^{s-1}}$  in the standard Kato-Ponce estimate).

**Lemma A.3** (Commutator estimates).

For any  $s \geq 0$ , and  $\partial_x f, g \in L^\infty(\mathbb{R}) \cap H^{s-1}(\mathbb{R})$ , one has

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

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

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

Let us now continue with the proof of Proposition 2.2. The system (1.1) is quasilinear. We prove below that it is Friedrichs-symmetrizable, under conditions (2.2). We display below the symmetrizer of the system, and compute the necessary energy estimates in Lemmata A.6 and A.7.

*Symmetrizer of the system.* Recall that (1.1) reads  $\partial_t U + A[U]\partial_x U = 0$ , with

$$A[U] \equiv \begin{pmatrix} u_1 & \frac{u_2 - u_1}{\varrho} & \frac{h_1}{\varrho} & \frac{h_2}{\varrho} \\ 0 & u_2 & 0 & h_2 \\ \frac{1}{\varrho} & 0 & u_1 & 0 \\ \frac{\gamma}{\varrho} & \delta + \gamma & 0 & u_2 \end{pmatrix}, \quad (\text{A.1})$$

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

$$S[U] \equiv \begin{pmatrix} \gamma & 0 & 0 & 0 \\ 0 & \gamma + \delta & 0 & u_2 - u_1 \\ 0 & 0 & \gamma h_1 & 0 \\ 0 & u_2 - u_1 & 0 & h_2 \end{pmatrix}. \quad (\text{A.2})$$

One can easily check that  $S[U]A[U] \equiv \Sigma[U]$  and  $S[U]$  are symmetric. More precisely, one has

$$\Sigma[U] \equiv \begin{pmatrix} \gamma u_1 & \frac{\gamma(u_2 - u_1)}{\varrho} & \frac{\gamma h_1}{\varrho} & \frac{\gamma h_2}{\varrho} \\ \frac{\gamma(u_2 - u_1)}{\varrho} & 2(\gamma + \delta)(2u_2 - u_1) & 0 & (\gamma + \delta)h_2 + u_2(u_2 - u_1) \\ \frac{\gamma h_1}{\varrho} & 0 & \gamma h_1 u_1 & 0 \\ \frac{\gamma h_2}{\varrho} & (\gamma + \delta)h_2 + u_2(u_2 - u_1) & 0 & h_2(2u_2 - u_1) \end{pmatrix}. \quad (\text{A.3})$$

One easily checks 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).

**Remark A.4.** Condition (2.2), which ensures the positive definiteness of our symmetrizer (and therefore the energy estimates proved below), is certainly too stringent. In particular, we do not recover the established domain of hyperbolicity in the limit  $\gamma \rightarrow 1$ , namely  $|u_2 - u_1|^2 < (1 + \delta^{-1})^2$ . It is in principle feasible to explicitly solve the eigenvalue/eigenvector problem for  $A[U]$ , determine the domain of strict hyperbolicity and consequently construct a corresponding symmetrizer with the help of spectral projectors. However, the calculations become extremely heavy, thus we choose to stick with the simpler symmetrizer in (A.2), acknowledging the lack of optimality. We let the reader refer to [1, 7, 31] for a more detailed analysis.

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

$$\begin{aligned} E^s(U) &\equiv (S[U]\Lambda^s U, \Lambda^s U) \\ &= \gamma |\zeta_1|_{H^s}^2 + (\gamma + \delta) |\zeta_2|_{H^s}^2 + \gamma \int_{\mathbb{R}} \underline{h}_1 |\Lambda^s u_1|^2 + \int_{\mathbb{R}} \underline{h}_2 |\Lambda^s u_2|^2 + 2 \int_{\mathbb{R}} (u_2 - u_1) \{ \Lambda^s u_2 \} \{ \Lambda^s \zeta_2 \}, \end{aligned} \quad (\text{A.4})$$

with  $\underline{h}_1 \equiv 1 + \varrho\zeta_1 - \zeta_2$  and  $\underline{h}_2 \equiv \delta^{-1} + \zeta_2$ .

We precise below 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:

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

**Lemma A.5.** *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  $C_1 = C(|\underline{h}_1|_{L^\infty}, |\underline{h}_2|_{L^\infty}, \delta_{\max}) > 0$  such that*

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

*Proof.* The fact that  $E^s(U) \leq C_1 |U|_{X^s}^2$  is a simple consequence of Cauchy-Schwarz inequality, applied to (A.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, one has

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

and the result is now clear. Lemma A.5 is proved.  $\square$

We now highlight energy estimates concerning the linearized system from (1.1), namely

$$\partial_t U + A[\underline{U}] \partial_x U = \mathcal{R}, \quad (\text{A.5})$$

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

**Lemma A.6** ( $L^2$  energy estimate). *Set  $T, M > 0$ . Let  $U \in L^\infty([0, T]; X^0)$  satisfy (A.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*

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

*Then there exists  $C_0 \equiv C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$  such that*

$$\forall t \in [0, T], \quad E^0(U)(t) \leq e^{C_0 M \varrho^{-1} t} E^0(U|_{t=0}) + C_0 \int_0^t e^{C_0 M \varrho^{-1} (t-t')} |\mathcal{R}(t', \cdot)|_{X^s} dt'. \quad (\text{A.6})$$

*Proof.* Let us consider the  $L^2$ -inner product of (A.5) and  $S[\underline{U}]U$ :

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

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

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} E^0(U) &= \frac{1}{2} (U, [\partial_t, S[\underline{U}]]U) - (\Sigma[\underline{U}] \partial_x U, U) + (\mathcal{R}, S[\underline{U}]U) \\ &= \frac{1}{2} (U, [\partial_t, S[\underline{U}]]U) + \frac{1}{2} ([\partial_x, \Sigma[\underline{U}]]U, U) + (\mathcal{R}, S[\underline{U}]U). \end{aligned} \quad (\text{A.7})$$

We now estimate each of the terms in the right-hand side of (A.7).

*Estimate of  $(U, [\partial_t, S[\underline{U}]]U)$ .* One has  $(U, [\partial_t, S[\underline{U}]]U) = (U, dS[\partial_t \underline{U}]U)$ , with

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

Using Cauchy-Schwarz inequality, and Lemma A.5, one has straightforwardly

$$|(U, [\partial_t, S[\underline{U}]]U)| \leq C_0 |\partial_t \underline{U}|_{L^\infty} C_2^{-1} |U|_{X^0}^2 \leq C_0 M \varrho^{-1} E^0(U), \quad (\text{A.8})$$

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

Estimate of  $([\partial_x, \Sigma[\underline{U}]]U, U)$ . One has  $([\partial_x, \Sigma[\underline{U}]]U, U) = (U, d\Sigma[\underline{U}]U)$  with

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

As above, Cauchy-Schwarz inequality and Lemmata A.1 and A.5 yield

$$|(\Sigma[\underline{U}]\partial_x U, U)| \leq C_0 M \varrho^{-1} E^0(U), \quad (\text{A.9})$$

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

Estimate of  $(\mathcal{R}, S[\underline{U}]U)$ . By Cauchy-Schwarz inequality and Lemmata A.1 and A.5,

$$|(\mathcal{R}, S[\underline{U}]U)| \leq C_0 |U|_{X^s} |\mathcal{R}|_{X^s} \leq C'_0 E^s(U)^{1/2} |\mathcal{R}|_{X^s}, \quad (\text{A.10})$$

with  $C_0, C'_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ .

Estimate (A.6) is now a consequence of Gronwall-Bihari's inequality applied to the differential inequality obtained when plugging (A.8), (A.9), (A.10) into (A.7).  $\square$

**Lemma A.7** ( $H^s$  energy estimate). *Set  $M, T > 0$  and  $s \geq s_0 + 1$ ,  $s_0 > 1/2$ . Let  $U \in L^\infty([0, T]; X^s)$  satisfy (A.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*

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

Then there exists  $C_0 \equiv C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$  such that

$$\forall t \in [0, T], \quad E^s(U)(t) \leq e^{C_0 M \varrho^{-1} t} E^s(U|_{t=0}) + C_0 \int_0^t e^{C_0 M \varrho^{-1} (t-t')} |\mathcal{R}(t', \cdot)|_{X^s} dt'. \quad (\text{A.11})$$

*Proof.* As previously, we deduce from (A.5) the identity

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

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

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} E^s(U) &= \frac{1}{2} (\Lambda^s U, [\partial_t, S[\underline{U}]]\Lambda^s U) - (S[\underline{U}]\Lambda^s A[\underline{U}]\partial_x U, \Lambda^s U) + (\Lambda^s \mathcal{R}, S[\underline{U}]\Lambda^s U) \\ &= \frac{1}{2} (\Lambda^s U, [\partial_t, S[\underline{U}]]\Lambda^s U) + \frac{1}{2} ([\partial_x, \Sigma[\underline{U}]]\Lambda^s U, \Lambda^s U) + (\Lambda^s \mathcal{R}, S[\underline{U}]\Lambda^s U) \\ &\quad - (S[\underline{U}][\Lambda^s, A[\underline{U}]]\partial_x U, \Lambda^s U). \end{aligned} \quad (\text{A.12})$$

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

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

Similarly, one has

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

The remaining term is estimated as follows. Using the commutator estimate in Lemma A.3, one has

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

with  $C_0 = C(M, h_0^{-1}, \delta_{\min}^{-1}, \delta_{\max})$ . Altogether, one deduces from (A.12)

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

Estimate (A.11) is now a consequence of Gronwall-Bihari's inequality, and the Lemma is proved.  $\square$

The well-posedness of system (1.1) is now a consequence of the energy estimates of Lemmata A.6 and A.7, following the standard strategy (we let the reader refer to standard textbooks, *e.g.* [2,28,32], for more details). More precisely, the solution of the nonlinear problem (1.1),  $U$ , is obtained as the limit of an iterative scheme:

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

The restriction on the timescale  $t \in [0, T_\varrho]$  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 time domain independent of  $n$ . The desired estimate on  $|U|_{X^s}$  follows directly from Lemma A.7 with  $\underline{U} = U$  and  $R \equiv 0$ , and the corresponding estimate on  $|\partial_t U|_{X^s}$  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  $t \rightarrow T_{\max}$  if  $T_{\max} < \infty$  follows from standard continuation arguments. This concludes the proof of Proposition 2.2.

**Acknowledgements.** The author is grateful to Christophe Cheverry, Jean-François Coulombel and Frédéric Rousset for helpful advice and stimulating discussions. This work received the support of the Agence Nationale de la Recherche (project ANR BLANC SIMI 1 2013 DYFICOLTI).

## References

- [1] R. Abgrall and S. Karni. Two-layer shallow water system: a relaxation approach. *SIAM J. Sci. Comput.*, 31(3):1603–1627, 2009.
- [2] S. Alinhac and P. Gérard. *Opérateurs pseudo-différentiels et théorème de Nash-Moser*. Savoirs Actuels, 1991.
- [3] R. Barros, S. L. Gavriluk, and V. M. Teshukov. Dispersive nonlinear waves in two-layer flows with free surface. I. Model derivation and general properties. *Stud. Appl. Math.*, 119(3):191–211, 2007.
- [4] T. B. Benjamin. Internal waves of finite amplitude and permanent form. *J. Fluid Mech.*, 25(2):241–270, 1966.
- [5] J. L. Bona, T. Colin, and D. Lannes. Long wave approximations for water waves. *Arch. Ration. Mech. Anal.*, 178(3):373–410, 2005.
- [6] J. L. Bona, D. Lannes, and J.-C. Saut. Asymptotic models for internal waves. *J. Math. Pures Appl. (9)*, 89(6):538–566, 2008.
- [7] M. J. Castro-Díaz, E. D. Fernández-Nieto, J. M. González-Vida, and C. Parés-Madroñal. Numerical treatment of the loss of hyperbolicity of the two-layer shallow-water system. *J. Sci. Comput.*, 48(1-3):16–40, 2011.
- [8] W. Choi and R. Camassa. Weakly nonlinear internal waves in a two-fluid system. *J. Fluid Mech.*, 313:83–103, 1996.
- [9] W. Craig, P. Guyenne, and H. Kalisch. Hamiltonian long-wave expansions for free surfaces and interfaces. *Comm. Pure Appl. Math.*, 58(12):1587–1641, 2005.
- [10] W. Craig, P. Guyenne, and C. Sulem. Coupling between internal and surface waves. *Natural Hazards*, 57(3):617–642, 2010.
- [11] B. de Saint-Venant. 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.
- [12] V. Duchêne. 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.

- [13] V. Duchêne. Boussinesq/Boussinesq systems for internal waves with a free surface, and the KdV approximation. *M2AN Math. Model. Numer. Anal.*, 46:145–185, 2011.
- [14] V. Duchêne. Decoupled and unidirectional asymptotic models for the propagation of internal waves. *M3AS:Math. Models Methods Appl. Sci.*, 24(01), 2014.
- [15] A. E. Gill. *Atmosphere-ocean dynamics*, volume 30 of *International geophysics series*. Academic Press, 1982.
- [16] R. Grimshaw, E. Pelinovsky, and O. Poloukhina. 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.
- [17] P. Guyenne, D. Lannes, and J.-C. Saut. Well-posedness of the Cauchy problem for models of large amplitude internal waves. *Nonlinearity*, 23(2):237–275, 2010.
- [18] K. R. Helfrich and W. K. Melville. Long nonlinear internal waves. In volume 38 of *Annu. Rev. Fluid Mech.*, pages 395–425, 2006.
- [19] C. R. Jackson. An atlas of internal solitary-like waves and their properties. Accessible at url [http://www.internalwaveatlas.com/Atlas2\\_index.html](http://www.internalwaveatlas.com/Atlas2_index.html).
- [20] T. Kato. The Cauchy problem for quasi-linear symmetric hyperbolic systems. *Arch. Rational Mech. Anal.*, 58(3):181–205, 1975.
- [21] T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, 1995. Reprint of the 1980 edition.
- [22] T. Kato and G. Ponce. Commutator estimates and the Euler and Navier-Stokes equations. *Comm. Pure Appl. Math.*, 41(7):891–907, 1988.
- [23] D. Lannes. Secular growth estimates for hyperbolic systems. *J. Differential Equations*, 190(2):466–503, 2003.
- [24] D. Lannes. Sharp estimates for pseudo-differential operators with symbols of limited smoothness and commutators. *J. Funct. Anal.*, 232(2):495–539, 2006.
- [25] D. Lannes. *The water waves problem*, volume 188 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2013.
- [26] D. Leonardi. *Internal and Surface Waves in a Two-Layer Fluid*. PhD thesis, University of Illinois, 2011.
- [27] R. R. Long. On the Boussinesq approximation and its role in the theory of internal waves. *Tellus*, 17(1):46–52, 1965.
- [28] G. Métivier. *Para-differential calculus and applications to the Cauchy problem for nonlinear systems*, volume 5 of *Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series*, 2008.
- [29] G. Schneider and C. E. Wayne. The long-wave limit for the water wave problem. I. The case of zero surface tension. *Comm. Pure Appl. Math.*, 53(12):1475–1535, 2000.
- [30] L. F. Shampine and M. W. Reichelt. The MATLAB ODE suite. *SIAM J. Sci. Comput.*, 18(1):1–22, 1997.
- [31] A. L. Stewart and P. J. Dellar. Multilayer shallow water equations with complete coriolis force. part 3. hyperbolicity and stability under shear. *J. Fluid Mech.*, 723:289–317, 5 2013.
- [32] M. E. Taylor. *Partial differential equations. III Nonlinear equations*, volume 117 of *Applied Mathematical Sciences*. Springer-Verlag, 1997.
- [33] L. N. Trefethen. *Spectral methods in MATLAB*, volume 10 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), 2000.