Reviewer#2.

The authors are most grateful for your comments. We have followed your suggestions and revised the manuscript accordingly in many places. Please, find our responses below.

I enjoyed reading this paper because the numerical simulations are of high quality, the experimental design is well devised, and the results yield interesting insight into the behavior of colliding nonlinear internal solitary waves with trapped cores. I have one suggestion for major revision although this won't require too much work, and some suggestions for clarification:

1) *My only suggestion for revision is that the authors remove the three-dimensional results and discussion of the mixing, dissipation, and energetics. I would only trust discussion of these if the authors demonstrated that they are truly resolved through discussion of grid resolution requirements for DNS, i.e. grid resolution via Kolmogorov scale. It is hard to imagine that the mixing is resolved given that the molecular diffusion is so small. In fact, unresolved two-dimensional simulations can lead to more mixing because the inverse energy cascade in two dimensions stretches density filaments and leads to more numerical mixing, even if the dissipation is actually lower in two dimensions (see Fringer and Street 2003; doi:10.1017/S0022112003006189, Arthur and Fringer, 2014; doi:10.1017/jfm.2014.641). An additional problem with discussion of the energetics in the paper is that the Reynolds number varies significantly for different runs. Arthur and Fringer (2014) showed that not accounting for Reynolds number effects can give a very different picture of the dynamics of breaking internal solitary waves on slopes. Such may be the case for the results in Figure 12, for which it is difficult to determine whether the behavior of the energy loss is due to alpha effects or Reynolds number effects. It may be that the two-dimensional simulations represent the energetics to a reasonable degree, as in many studies of internal wave energetics, although I would not necessarily trust the arguments concerning the mixing. Either way, I suggest that the authors discuss the three-dimensional effects and associated energetics in a different paper*.

Answer. We have revised the text accordingly:
(i) Results of 3D simulation were excluded from this manuscript;
(ii) The effect of Re_m was discussed in text;

p. 9 l. 8 "The absence of complete self-similarity on the Reynolds and Schmidt numbers also means that the Euler equations do not describe the wave interaction processes in deep water even for the range of stable waves. As shown in Table 1, the parameter Re_m varies in Series A-C several times for waves of the same dimensionless amplitude α. The incomplete similarity scaling following Barenblatt (1996) results in relation: $\Delta E_{loss} = \Psi(\alpha)\mathrm{Re}_m^m \mathrm{Sc}^n$, where Ψ is function of α, m and n are exponents. However, this rescaling also did not result in universal dependence. We conclude that it is due to the different mechanisms governing collision process in ranges I-III: nonlinear wave interaction, collapse of collided trapped masses and instability. Another factor influencing the interaction may be the diffusivity effect (Deepwell and Stastna, 2016), which is described by the Schmidt number. However, in these experiments, the Schmidt number was large and constant. "
(iii) Text was added on limitations using 2D setting:

p. 10 l.8 “Notice, however, that the destruction of the KH billows is essentially three-dimensional process, therefore, 3D high-resolution simulation is necessary to predict turbulence development (Arthur and Fringer,2014, Deepwell and Stasna, 2016). This is the subject of a separate study, whereas the interaction of the colliding waves as a whole can be described in 2D setting.”

2) *Please discuss how you chose the grid resolution for the two-dimensional simulations.*

Answer. We carried out doubling-grid tests to verify that chosen grid adequately described flow fields. The comparison for wave A13 is show in Figs. 1 and 2. The text was added accordingly.

p. 4 l. 25 “Most of the runs were performed in a two-dimensional setting with a grid resolution of 3000×400 (length and height, respectively), whereas several runs for waves A9-A13 were also carried out with a grid resolution of 6000×800 (length and height, respectively) to verify effect of grid resolution on the wave interaction and to make the fine structure clearer. Comparison of the baseline and doubled grid resolution showed the equivalence of the calculated fields, with the exception of wave A13 for which 6000×800 resolution was used.”

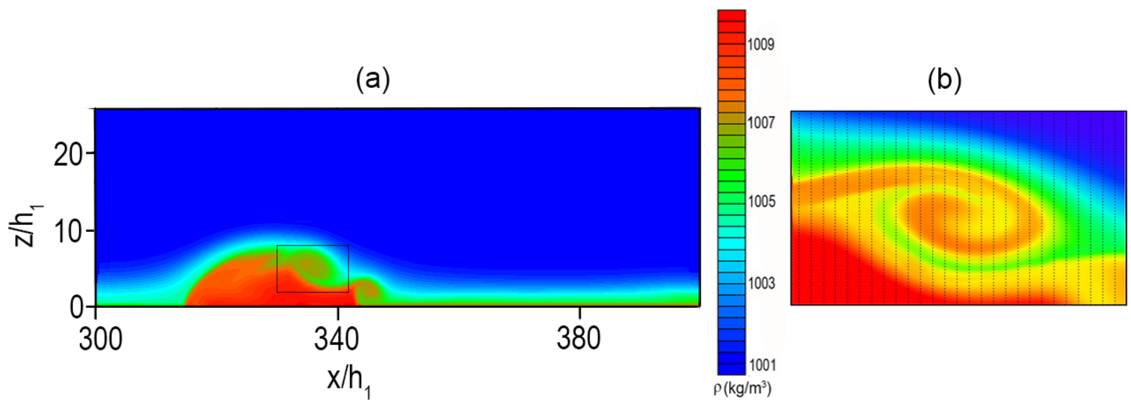

Fig. A1 Snapshot of the density field for case (A13;A13) at $\tau=175$ and Sc=1000 for grid resolution 3000x400 (a) and extended snapshot of KH billow with grid points (b).

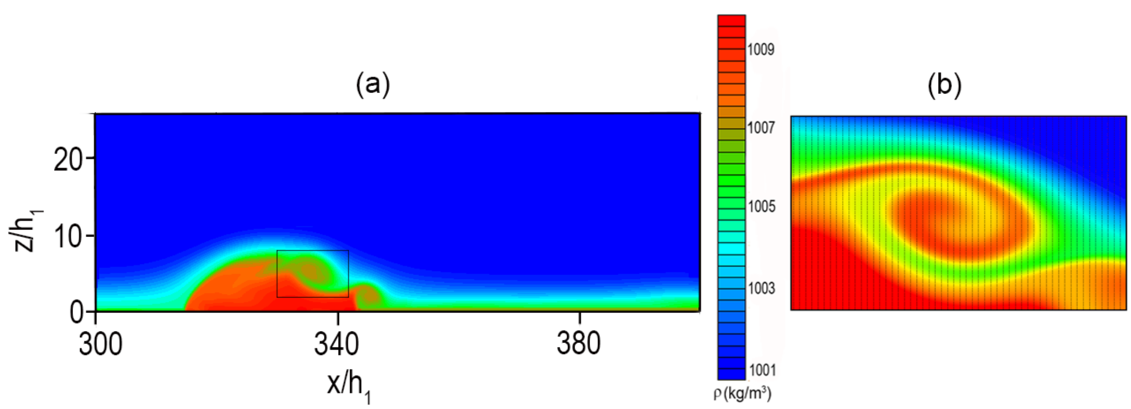

Fig. A2 Snapshot of the density field for case (A13;A13) at $\tau=175$ and Sc=1000 for grid resolution 6000x800 (a) and extended snapshot of KH billow with grid points (b).

3) The Richardson number should be defined as Ri_m=g'h/(U_m 2) so that it is consistent with the way the other nondimensional parameters are defined, i.e. in terms of the independent parameters following the Buckingham Pi theorem.

Answer. While there are various estimates of Froude, Richardson and Reynolds numbers, we found (Maderich etal., 2015) that definitions (5)-(7) allow one to adequately classify the state of colliding waves of large amplitude using local characteristics such as minimum Richardson number or the ratio of maximal local velocity to the ISW phase velocity.

p.4 l. 5 "The important features of the ISWs can be described by dimensionless amplitude is $\alpha = a/h$, the Froude, Richardson and Reynolds numbers based on local characteristics of waves (Maderich et al.,2015)."

4) *It would be helpful if, on page 3, you discussed the general features of Series A-D, and included a brief description in another column in Table 1, i.e. a column indicated by "Comments" which, for series A would state, "No trapped cores". Also please indicate whether the waves were in regimes (i), (ii), or (iii) in Table 1.*

Answer. We added column with class of colliding waves. In accordance with the definition of the class of the ISW, waves with trapped cores belong to the classes (ii) and (iii).

5) *What is the justification for choosing such a small molecular diffusion?*

Answer. Text and figure were added to consider the impact of small diffusivity on the collision processes.

p.7 l. 8 "In the ocean and in the most of the laboratory experiments the Schmidt number is about 700-800. The used grid does not allow the whole range of inhomogeneities in salinity (density) to be resolved. Therefore, it is important to evaluate the effect of molecular diffusion of salinity on the dynamics of waves and to verify the possibility that diffusion can be neglected in the wave collision for large Sc. Two cases for large amplitude waves were considered (A9;A9) and (A13;A13). We performed runs for Sc=1; 10 and 1000. In the collision case (A9;A9) the behaviour of colliding waves are the same, whereas the difference between runs for Sc=1 and Sc=1000 was less than 1% of $\Delta\alpha/\alpha$ and $\Delta\theta$ values. The comparison of the density snapshots during collision in case (A13;A13) for different Schmidt numbers is shown in Fig. 9. Figure clearly depicts difference between structure of interacting waves for cases Sc=1 and Sc=10. The corresponding values of $\Delta\alpha/\alpha$ and $\Delta\theta$ differ by 5% and 0.6%, respectively. This was in agreement with the results by Deepwell and Stastna (2016), where it was shown essential effect of molecular diffusivity on the mass transport by mode-2 ISW in range $1 \leq \mathrm{Sc} < 20$. At the same time, the results of calculations at Sc=10 and Sc=1000 in Fig.9b and 9c practically coincide, which indicates that molecular diffusion may not be taken into account when studying the global properties of colliding waves. This conclusion agrees with (Terez and Knio, 1998) as they estimate that the value of Sc=100 was "sufficiently high for density diffusion to be ignored during simulation period" and the results of the Deepwell and Stastna (2016) simulation, according to which the mass transfer is virtually independent of Sc already at Sc>20. However, diffusion can be important for small scale mixing processes in tiny density structures (see e.g. Galaktionov et al., 2001) forming in result of instability and turbulent cascade processes (Deepwell and Stastna, 2015) and persisting over time in a wake behind moving bulge of trapped fluid (Terez and Knio, 1998). These subgrid scale structures in our simulations were smashed by numerical diffusion which did not affect larger scale due to use of second order total variation diminishing (TVD) scheme for advective terms in transport equation."

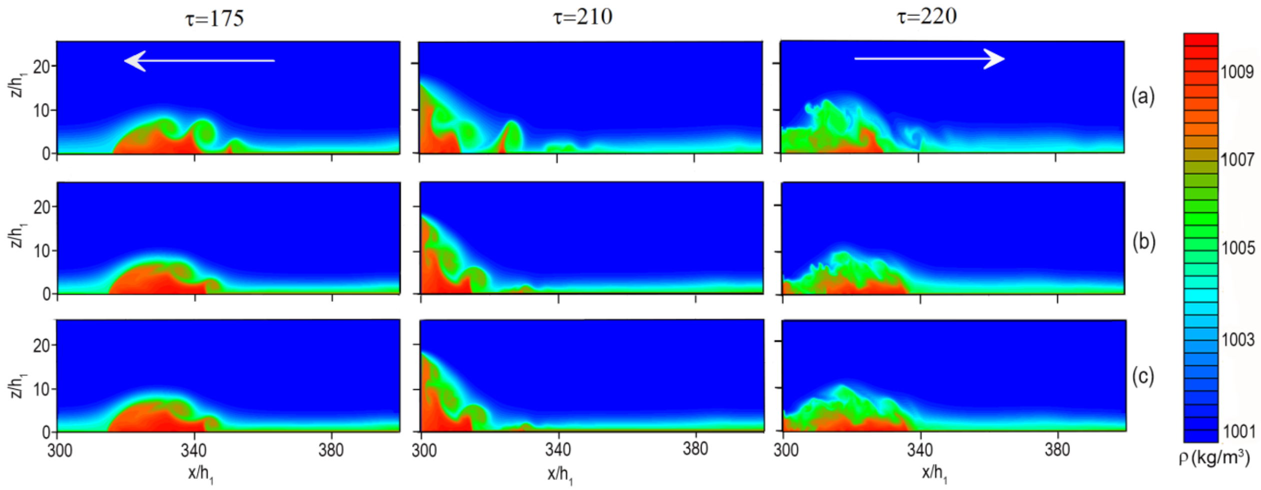

Figure 9. Comparison of the density snapshots during collision of ISWs in case (A13;A13) for different Schmidt numbers. (a) Sc=1. (b) Sc=10. (c) Sc=1000. The right half of the numerical flume is shown due to the symmetry of the interaction process.

6) *Page 4, Line 9: Please explain the meaning of and how you computed the phase shift* $\Delta\theta$ *, and how it is normalized by* τ_0 *.*

Answer. We estimated temporal phase shift by comparing trajectories of the wave crests with and without collision. This temporal phase shift was normalized to characteristic time

p.5 l. 9 "…whereas normalized to characteristic time τ_0 temporal phase shift $\Delta\theta$ is $\Delta\theta \sim \alpha$."
p.5 l. 10 "We estimated temporal phase shift by comparing trajectories of the wave crests with and without collision."

7) *Page 4, Line 10: Please explain how you expect Da/a and Dq to behave for limiting cases* ($\alpha \to 0$ *and* $\alpha \to \infty$)*. Why does* $\Delta\theta \to 4$ *as* $\alpha \to \infty$*?*

Answer. We added text with explanations.

p. 6 l. 12 "As shown in Fig. 4a, for stable waves of class (ii), the runup excess $\Delta\alpha/\alpha$ still almost linearly increases in the range $2.3 < \alpha < 4.6$, whereas the increase in the phase shift $\Delta\theta$ is substantially slowed down when $\alpha > 1$, and then $\Delta\theta$ tends towards a constant value $\alpha > 4$. The distributions of $\Delta\alpha/\alpha$ and $\Delta\theta$ in Fig. 4 for stable waves were approximated by linear and exponential dependencies, respectively, which were based on the weakly-nonlinear asymptotics $\Delta\alpha/\alpha \sim \alpha$ and $\Delta\theta \sim \alpha$ (Matsuno,1998) for small α and obtained in numerical experiments almost constant distribution of $\Delta\theta$ at large α. "

8) *On Page 5, Line 5, you state that the colliding waves pass through each other. Theory suggests that nonlinear waves exchange momentum by bouncing off each other, just like billiard balls (e.g. Fringer and Holm 2001; doi:10.1016/S0167-2789(00)00215-3).*

Answer. The collision of large amplitude ISWs with trapped cores is complicated process, which in theory has not yet been described in detail, in particular for waves of different amplitude. To avoid misinterpretation of results we changed “transmitted” waves to “outgoing” waves.

9) *Please do not include the regressions on Page 5, line 13, unless you can justify the functional relationships through scaling or other means.*

Answer. See answer to comment 7)

10) *Minor: a. I don’t understand the meaning of the sentence starting with “The waves of class (ii) …” on line 8 of page 1. b. Throughout: monotonous à monotonic.*

Answer. The text was changed accordingly.

p. 1 l. 9 “The colliding waves of class (ii) lose fluid trapped by the wave cores when a normalized on thickness of pycnocline amplitudes are in the range of approximately between 1 and 1.75.”