Journal of Advances in Modeling Earth Systems (JAMES)

RESEARCH ARTICLE

A low-storage Runge–Kutta framework for nonlinear free-surface ocean models

Simone Silvestri[1,2], **Jean-Michel Campin**[2], **Gregory Wagner**[2,3], **Navid C. Constantinou**[4,5], **Xin Kai Lee**[2,6], **and Raffaele Ferrari**[2,6]

[1]DIATI, Politecnico di Torino, Torino, Italy [2]Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA [3]Aeolus Labs, San Francisco, CA, USA [4]University of Melbourne, Parkville, VIC, Australia [5]Australian Research Council Center of Excellence for the Weather of the 21st Century, Australia [6]Center for Computational Science and Engineering, Massachusetts Institute of Technology

Key Points:

- Implicit and split-explicit formulations of the Runge-Kutta framework preserve conservation and consistency while ensuring stability.
- Modal analysis shows split-explicit Runge-Kutta methods require averaged barotropic transport velocity to maintain baroclinic stability.
- Third-order Runge–Kutta time-stepping reduces numerical mixing relative to Adams–Bashforth methods even when using a larger time-step size.

Correspondence to:
S. Silvestri,
silvestri.simone0@gmail.com

Citation:
Your citation here.

Abstract We present a comprehensive low-storage Runge–Kutta framework for nonlinear free-surface ocean models that is stable across the coupled barotropic–baroclinic modes and conserves tracers up to machine precision. We present two different formulations that depend on the time discretization of the coupled barotropic mode: *(i)* an implicit free surface discretization and *(ii)* a split–explicit free surface discretization. Our stability analysis reveals that the split–explicit implementation requires advecting the tracers with an averaged barotropic transport velocity, not only for tracer conservation reasons, but also to maintain baroclinic stability. By adapting exact variance budget diagnostics to multi-stage integrators, we are able to directly quantify the contribution of temporal discretization to numerical mixing. Across three idealized test cases spanning the linear to fully turbulent regimes, the three-stage Runge–Kutta formulation produces up to 50% lower numerical mixing than the Adams–Bashforth time discretization with the same spatial discretization. The framework is implemented in Oceananigans, providing the ocean modeling community with a practical path toward more accurate ocean simulations.

Plain Language Summary Ocean models simulate how currents, temperature, and salinity evolve over time, but the methods used to advance these simulations can introduce artificial blurring, referred to as “numerical mixing”, acting as an unwanted smoothing filter that gradually degrades the solution. We adopt a more accurate time-stepping approach for ocean models that produces less artificial blurring than traditional methods while keeping the simulation stable. In our tests, well-resolved physical processes dominated over numerical blurring, preserving temperature gradients and realistic eddy behavior. This improvement matters for climate projections, where even small amounts of numerical mixing accumulate over decades-long simulations, potentially biasing predictions of ocean circulation and heat transport.

1. Introduction

Ocean general circulation models must resolve processes spanning vastly different timescales, from external gravity waves propagating at hundreds of meters per second to eddy advection and internal waves occurring over days to weeks. This separation, typically of two to three orders of magnitude, has driven the development of split–explicit and implicit-explicit (IMEX) approaches that decouple the fast barotropic dynamics from the slower three-dimensional evolution (Killworth et al., 1991; Campin et al., 2004; Shchepetkin & McWilliams, 2005). The baroclinic component, governing currents, tracers, and stratification, advances on time steps limited by advective Courant–Friedrichs–Lewy (CFL) condition, while the barotropic mode employs either sub-cycling or implicit solvers to handle the fast gravity waves efficiently (e.g., Griffies, Pacanowski, Schmidt, and Balaji (2001)).

Despite the diversity in modeling approaches, baroclinic time-stepping in ocean models has converged toward a few low-order, single-stage methods that inherently require some form of dissipation for stability. Leapfrog schemes, typically used in atmospheric models (Klöwer et al., 2024), and employed in the ocean models ROMS (Shchepetkin & McWilliams, 2005), LICOM (Wang et al., 2021), NEMO (Madec et al., 2017), and FESOM (Danilov et al., 2017), offer nominally non-dissipative, second-order accurate integration. However, these leapfrog schemes are susceptible to computational mode instabilities and require filtering, such as the Asselin filter (Asselin, 1972) that is used in NEMO and LICOM, which reintroduces numerical dissipation. The MPAS-Ocean model (Ringler et al., 2013) uses a simpler one-stage staggered scheme that avoids computational modes but achieves only first-order accuracy. MOM6 (Adcroft et al., 2019) is an exception with its two-stage predictor-corrector approach based on a modified version of the second-order Runge–Kutta method; RK2 is unconditionally unstable for purely advective problems (see figure 1), requiring a lower-order implementation to provide the necessary stabilizing dissipation. The Adams–Bashforth method, which avoids time-staggering between state variables, remains the most prevalent choice, used in MITgcm (Marshall et al., 1997), ICON-O (Korn, 2017), Veros (Häfner et al., 2018), and Oceananigans (Wagner, Silvestri, et al., 2025). These ocean models mostly use quasi-first-order variants of the second-order Adams–Bashforth scheme; an exception is the third-order Adams–Bashforth scheme implemented in the MITgcm, which achieves higher accuracy at the same computational cost (Forget et al., 2015). However, the latter is rarely used since *(i)* it requires substantially more memory and *(ii)* it reduces the stability limit of diffusive processes, pairing quite poorly with high-order diffusive spatial discretizations such as WENO schemes (Silvestri et al., 2024).

Multi-stage schemes such as Runge–Kutta (RK) methods offer a promising alternative, providing higher-order accuracy and larger stable CFL numbers while also reducing implicit dissipation without requiring state variable staggering. Low-storage RK formulations, which require the same memory as a second-order Adams–Bashforth method, have been successfully implemented in atmospheric models (Wicker & Skamarock, 2002; Skamarock et al., 2008, 2012; Baldauf et al., 2011; Yatunin et al., 2025) and are increasingly being explored for ocean applications (Ishimwe et al., 2023; Lan et al., 2022). The low-storage property is particularly attractive for memory-constrained applications, as it allows high-order temporal accuracy without storing multiple time levels of the whole state vector. This is particularly important for GPU-based ocean models, such as Oceananigans, where memory leanness is a strict requirement (Silvestri et al., 2025). However, implementing RK schemes in free-surface ocean models presents unique challenges associated with multi-stage methods that extend beyond baroclinic time-stepping itself, and require careful coupling with barotropic free-surface evolution while maintaining essential numerical properties.

The above-mentioned challenges require satisfying three fundamental constraints simultaneously, each of which complicates the design of ocean-specific split-time-stepping schemes. First, *free surface reconciliation* requires consistency between the explicitly evolved free surface and the grid stretching diagnosed from the continuity equation in response to the free surface displacements; violations lead to spurious mass sources and unphysical vertical velocities (Campin et al., 2004). Second, *tracer conservation* must hold both globally (preserving the total tracer content) and locally (maintaining constant tracer fields constant), which becomes particularly challenging with moving vertical coordinates, where tracer and continuity discretizations must match exactly (Griffies et al., 2001). Third, *numerical stability* of the coupled barotropic–baroclinic system requires careful analysis of mode interactions, as improper coupling can destabilize either fast or slow modes (Demange et al., 2019).

Several recent studies have advanced RK implementations for ocean models. Ishimwe et al. (2023) developed a robust split–explicit RK3 scheme paired with a Discontinuous Galerkin (DG) spatial discretization, using a three-stage RK scheme for the primitive equations coupled with a two-stage RK scheme for the barotropic mode. They show that their method is stable and tracer-conserving. In contrast, Lan et al. (2022) implemented the classical three-stage strong-stability-preserving (SSP) Runge–Kutta formulation for both the three-dimensional evolution and the free-surface evolution in the MPAS-Ocean framework. Most recently, NEMO has transitioned from its Modified Leap Frog scheme to a third-order Runge–Kutta method in version 5.0, citing improved stability and computational efficiency (NEMO System Team, 2024). Building on these foundations, our work revisits the Runge–Kutta framework to simultaneously address all three constraints mentioned above: free-surface reconciliation, tracer conservation, and stability, within a low-storage RK formulation paired with both an implicit and a split–explicit free-surface discretization in a finite volume ocean model.

We focus on comparing this new framework with the quasi-second-order Adams–Bashforth (AB2) method for two reasons: AB2 is widely used across ocean models, and it is the only current time discretization for the Oceananigans hydrostatic model (Silvestri et al., 2025), the ocean model we aim to improve. AB2 offers significant computational advantages: it requires only one tendency evaluation per time step and minimal memory overhead by storing just two time levels; moreover, it computes the entire state vector in a single evaluation, avoiding the time-staggering between variables necessary in leapfrog schemes. However, AB2 has a rather small stability range and introduces implicit numerical dissipation, as indicated by amplification factors that are significantly less than one across a wide range of CFL numbers (Figure 1). While the impact of spatial discretization on numerical mixing, the spurious diffusion of tracers introduced by discretization errors, has been extensively studied, where spurious diapycnal mixing from advection schemes have been shown to rival or exceed explicitly parameterized physical mixing (Klingbeil et al., 2014; Hill et al., 2012), the contribution of time discretization to numerical mixing has received comparatively less attention. The dissipative nature of AB2 suggests that time-stepping schemes may also contribute substantially to spurious mixing in realistic ocean configurations and degrade model fidelity by affecting ocean stratification over long integrations (Ilicak et al., 2012).

To quantify the contribution of time-stepping to numerical mixing, robust diagnostic tools have been developed. Klingbeil et al. (2014) introduced methods for analyzing the relative contributions of physical and numerical mixing by analyzing tracer variance decay. Hill et al. (2012) used the spreading of virtual tracer release experiments. Ilicak et al. (2012) introduced the reference potential energy (RPE) framework, an exact, globally integrated measure of numerical mixing, and subsequently extended this approach to quantify the spatial distribution of numerical mixing (Ilicak, 2016). More recently, Banerjee, Danilov, Klingbeil, and Campin (2024) extended the variance budget method to provide diagnostics that are exact locally in both space and time, a property essential for isolating the contribution of time discretization to numerical mixing. However, most existing diagnostics have been developed and validated for single-step methods rather than multi-stage RK schemes, a gap this study aims to fill.

In this study, we develop a low-storage RK time-stepping scheme and couple it scheme with both the implicit (RK-IM) and the split–explicit (RK-SE) free-surface time discretization, demonstrating how each approach maintains free-surface reconciliation and tracer conservation through careful time-discretization design. We provide a detailed linear stability analysis of the coupled barotropic–baroclinic system, extending the modal decomposition approach of Demange et al. (2019) to multi-stage RK schemes and revealing the importance of choosing the right coupling velocities to maintain stability in both the barotropic and baroclinic modes. We adapt the exact variance budget methodology to quantify numerical mixing in multi-stage RK schemes and the multi-level AB2 time discretization. Finally, we demonstrate the advantages of our RK schemes in three idealized test cases: topographic generation of an internal tide, coastal baroclinic adjustment, and equilibrated channel flow. These three cases stress different aspects of ocean dynamics and reveal substantial reductions in numerical mixing compared to AB2 methods.

The remainder of this paper is organized as follows. In section 2, we introduce the low-storage RK formulation and compare its stability and dissipation properties with AB2 in the context of a generic ODE (section 2). We detail the implementation of the primitive equations, addressing free-surface consistency and tracer conservation for both implicit and split–explicit treatments in section 3, followed by a linear stability analysis of the coupled barotropic–baroclinic system in section 4. Finally, we demonstrate the advantages of the proposed schemes through three idealized test cases that reveal a substantial reduction in numerical mixing compared to AB2 in section 5 and conclude in section 6.

2. A low-storage Runge–Kutta time discretization

Our goal is to replace the quasi-second-order Adams–Bashforth scheme currently used in Oceananigans' hydrostatic model (Silvestri et al., 2025; Wagner, Silvestri, et al., 2025) with a Runge–Kutta discretization that improves stability and reduces numerical dissipation while maintaining low memory requirements. We wish to apply the Runge–Kutta discretization to the the baroclinic (slow) component of the governing equations and pair it with a separate time discretization for the barotropic (fast) component, enabling flexible comparisons across different approaches.

For a simple ODE of the form,

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = F(\varphi)\,, \tag{1}$$

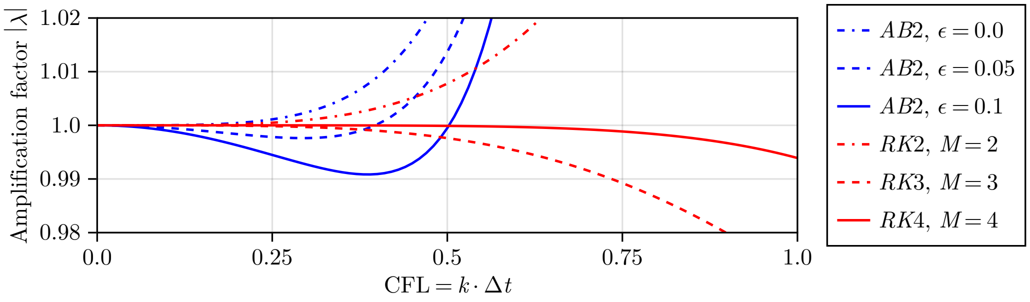

Figure 1: Amplification factors as a function of the CFL number ($k\Delta t$) for different time discretizations. $AB2$ stands for Quasi-Adams–Bashforth 2nd order and RK stands for a multi-stage Runge–Kutta, with M the number of stages.

the quasi-second-order Adams–Bashforth (AB2) implementation reads,

$$\varphi^{n+1} = \varphi^n + \Delta t \left[\left(\frac{3}{2} + \epsilon \right) F(\varphi^n) - \left(\frac{1}{2} + \epsilon \right) F(\varphi^{n-1}) \right] , \tag{2}$$

where typically $\epsilon = 0.1$. A commonly used M-stage Runge-Kutta formulation for the same ODE (1) reads,

$$\begin{cases} \varphi^{n,0} &= \varphi^n \, , \\ \varphi^{n,m+1} &= \varphi^n + \dfrac{\Delta t}{M - m} F\left(\varphi^{n,m}\right) \quad \text{for} \;\; m = 0, \ldots, M-1 \, , \\ \varphi^{n+1} &= \varphi^{n,M} \, . \end{cases} \tag{3}$$

The second superscript m refers to the RK stage within the n-th time step, so that $\varphi^{n,0} \equiv \varphi^n$ and $\varphi^{n,M} \equiv \varphi^{n+1}$. With a slight abuse of notation, we drop the leading n in flowing prose and write φ^m for $\varphi^{n,m}$ throughout the rest of the paper, retaining φ^n and φ^{n+1} only for the boundary values at the start and end of the time step; the algorithm summaries in section 3 carry the full (n, m) superscript so they can be read out of context. This formulation was originally developed by Wicker and Skamarock (2002) for split–explicit atmospheric models, where its simplicity and low memory requirements proved advantageous for time-splitting schemes. Ishimwe et al. (2023) subsequently adapted it for ocean applications, using two RK stages for the barotropic component and three stages for the baroclinic component. An alternative is the classical three-stage Runge–Kutta used by Lan et al. (2022), which provides strong stability-preserving (SSP) properties but requires a more complex formulation that complicates barotropic–baroclinic coupling.

We adopt the formulation by Wicker and Skamarock (2002) for its simplicity of implementation and straightforward coupling with barotropic solvers. While it lacks the SSP property, it achieves the same numerical properties as SSP schemes for linear systems.

To compare the dissipation characteristics of RK and AB2, we analyze their stability for a linear tracer advection equation: $\partial_t \varphi + \partial_x \varphi = 0$. Projecting onto Fourier modes, the equation takes the form: $\partial_t \hat{\varphi} = -ik\hat{\varphi}$ where $\hat{\varphi}$ is the Fourier mode for wavenukmber k. We can rewrite the above time discretization as

$$\hat{\varphi}^{n+1} = \hat{\varphi}^n \left[1 + \sum_{m=1}^{M} \frac{(-i)^m k^m \Delta t^m}{m!} \right] . \tag{4}$$

We also rewrite the AB2 time discretization of the linear tracer advection equation as

$$\hat{\varphi}^{n+1} = \hat{\varphi}^n \left[1 - ik\Delta t \left(\frac{3}{2} + \epsilon \right) + ik\Delta t \left(\frac{1}{2} + \epsilon \right) \frac{\hat{\varphi}^{n-1}}{\hat{\varphi}^n} \right] . \tag{5}$$

Then, the amplification factor λ for the M-stage RK discretization is

$$|\lambda| \overset{\text{def}}{=} \left| \frac{\hat{\varphi}^{n+1}}{\hat{\varphi}^{n}} \right| = \left| 1 + \sum_{m=1}^{M} \frac{(-i)^m k^m \Delta t^m}{m!} \right| , \tag{6}$$

while for the AB2, we obtain

$$|\lambda| = \frac{1}{2} \left| 1 - \left(\frac{3}{2} + \epsilon\right) ik\Delta t + \sqrt{1 - i(1-2\epsilon)k\Delta t - \left(\frac{3}{2}+\epsilon\right)^2 k^2 \Delta t^2} \right| . \tag{7}$$

A time discretization is stable only for $|\lambda| \leq 1$, with values closer to one being more desirable because of the lower implied numerical dissipation, since each Fourier mode is advected independently and linearly without growth or decay.

Figure 1 compares the amplification factors (6) and (7) for different values of ϵ and M. Both AB2 with $\epsilon = 0$ and RK2 with $M = 2$ are linearly unstable ($|\lambda| > 1$) for purely advective systems and require explicit dissipation for stability. With $M > 2$, the RK discretization remains stable over a much larger CFL range than AB2 and introduces substantially less implicit dissipation, a key advantage for long-term integrations where numerical mixing accumulates.

It should be noted that the RK discretization requires more right-hand-side evaluations per time step than AB2, thereby increasing computational cost. To ensure fair comparisons, we use larger time steps for RK than for AB2 throughout the remainder of this paper, matching computational cost rather than CFL number. The following section details how we apply this RK formulation to the primitive equations.

3. Time discretization of the Primitive equations

We set out to derive an RK discretization of the primitive equations. Readers primarily interested in the final time-stepping procedures may skip directly to the flowcharts Algorithm 1 and Algorithm 2 at the end of this section.

Throughout this section we follow the indexing convention introduced in section 2: φ^n denotes a quantity at time step n (equivalently $\varphi^{n,0}$, the start of the time step), $\varphi^{n,m}$ denotes the value at the m-th RK stage of time step n, and $\varphi^{n+1} \equiv \varphi^{n,M}$ denotes the end of the time step. In text and equations we suppress the leading n and write φ^m for $\varphi^{n,m}$, since the substep equations always refer to stages within the same time step; the symbols φ^n and φ^{n+1} continue to denote the boundary values at $m = 0$ and $m = M$. Algorithms 1 and 2 carry the full (n, m) superscript on every variable so they can be read out of context. We use a superscript asterisk to denote a *predictor* quantity (e.g., $\mathbf{u}^*$, $\mathbf{U}^*$) computed before the free-surface correction, and a superscript dagger to denote a quantity *corrected* so as to enforce local tracer conservation (e.g., $\mathbf{u}^\dagger$, $\mathbf{U}^\dagger$, $\omega^\dagger$).

We use a z-star vertical coordinate formulation (Adcroft & Campin, 2004), with the vertical coordinate z given by

$$z = r\frac{H+\eta}{H} + \eta \, . \tag{8}$$

Here, H is the water column height at rest, η is the free surface elevation as measured from the rest state of the ocean ($z = 0$), and r is a reference (fixed) geopotential vertical coordinate with $r = 0$ at $z = \eta$ and $r = -H$ at $z = -H$. We note that in the literature this vertical coordinate is often denoted z^*; we use r throughout to keep the integration variable distinct from the physical height z. The derivative of the vertical coordinate with respect to r is called “specific thickness” and usually denoted with σ:

$$\sigma = \frac{\partial z}{\partial r} = \frac{H+\eta}{H} \, . \tag{9}$$

The primitive equations in r coordinates are derived in many manuals and papers (e.g., Adcroft & Campin, 2004); for completeness, we reproduce the derivation in Appendix A. They read:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{G}_u - g\boldsymbol{\nabla}\eta \,, \tag{10}$$

$$\frac{\partial p}{\partial r} = \sigma b \,, \tag{11}$$

$$\frac{\partial \sigma}{\partial t} = -\boldsymbol{\nabla}\cdot(\sigma\mathbf{u}) - \frac{\partial \omega}{\partial r} \,, \tag{12}$$

$$\frac{\partial \sigma C}{\partial t} = -\boldsymbol{\nabla}\cdot(\sigma\mathbf{u}C) - \frac{\partial(\omega C)}{\partial r} \,, \tag{13}$$

where g is the gravitational acceleration, $\boldsymbol{\nabla}$ is the horizontal gradient operator $(\hat{\mathbf{x}}\,\partial_x + \hat{\mathbf{y}}\,\partial_y)$ at constant r, $\mathbf{u}$ is the horizontal velocity field $u\hat{\mathbf{x}}+v\hat{\mathbf{y}}$, ω is the vertical velocity across r surfaces, C indicates a (passive or active) tracer, p is the hydrostatic pressure, and b is the buoyancy (which depends on the temperature T, salinity S, and depth z). Finally, $\mathbf{G}_u$ includes all the "slow" baroclinic tendency terms (advection, Coriolis, hydrostatic pressure gradient, vertical mixing, and boundary conditions). The above equations are complemented by the evolution equation for the free surface, which, in the absence of precipitation or evaporation, reads

$$\frac{\partial \eta}{\partial t} = -\boldsymbol{\nabla}\cdot\int_{-H}^{0}\sigma\mathbf{u}\,\mathrm{d}r \,. \tag{14}$$

The free surface must also be consistent with equation (8), which implies,

$$\eta = \underbrace{\int_{-H}^{0}\sigma\,\mathrm{d}r}_{\overset{\text{def}}{=}\mathcal{H}} - H \,, \tag{15}$$

where we denoted the column thickness with $\mathcal{H} = H + \eta$.

We require our time discretization scheme to satisfy the following numerical properties:

1 The free surface is evolved consistently with the grid thickness given by equation (15). This condition is typically referred to as free surface reconciliation.

2) Tracers are conserved both globally (the integrated value of a tracer field remains constant in the absence of external fluxes) and locally (an initially constant tracer field remains constant in the absence of external fluxes).

3) The discretization is robust and stable for sufficiently large time steps.

3.1. Condition 1: free surface consistency

The evolution equation for the grid is the continuity equation (12). In our model, this equation is used to diagnose vertical velocity (ω); the grid is not explicitly evolved, as in models that use more sophisticated vertical coordinates (e.g., Adcroft et al., 2019), but is diagnosed from equation (8). This ensures that (15) is always satisfied.

3.2. Condition 2: tracer conservation

Tracer conservation is crucial to avoid model drift and unphysical behavior of temperature and salinity in long-time integrations (Griffies et al., 2001). In ocean models with a moving grid, achieving conservation requires careful design of the time discretization. The evolution equation for a constant tracer simplifies to the continuity equation (12). Therefore, to ensure local tracer conservation, the time discretization of the grid evolution necessarily needs to match the time discretization of the tracer equation. Since we use a z-star formulation, the grid evolves with the free surface

$$\frac{\partial \sigma}{\partial t} = \frac{1}{H}\frac{\partial \eta}{\partial t} = -\frac{1}{H}\boldsymbol{\nabla}\cdot\int_{-H}^{0}\sigma\mathbf{u}\,\mathrm{d}r \,. \tag{16}$$

If we use an explicit free surface discretization that matches the tracer equation, tracer conservation is automatically guaranteed. However, the characteristic speed of the gravity waves associated with the free surface is much faster than the typical advection velocity. Thus, using an explicit free-surface time discretization that pairs with the tracer evolution imposes an extreme CFL restriction that renders the scheme useless for practical purposes. This brings us to condition 3.

3.3. Condition 3: algorithm robustness

To ensure numerical stability and overcome the stringent CFL condition imposed by gravity waves, two approaches are available for the free surface evolution: an implicit discretization (Dukowicz & Smith, 1994) or a split–explicit sub-cycling approach, where the barotropic mode is advanced with a much smaller timestep within each baroclinic time step (Killworth et al., 1991; Shchepetkin & McWilliams, 2005). We implement both approaches, coupling the free surface and the three-dimensional evolution at each RK substep. A coupling at each substep corresponds to an implicit-explicit (IMEX) RK scheme with the same coefficients for the explicit and implicit component. A von Neumann stability analysis was performed in the appendix of Kar (2006), showing that this scheme is first order in the implicit component (in this case, the barotropic mode). More sophisticated implicit-explicit RK schemes (Conde et al., 2017; Rokhzadi et al., 2018) could achieve higher implicit order, but here we prioritize simplicity. The globally stiffly accurate coupling ensures stability while providing an effective improvement over the diffusive AB2 time discretization. We tested several alternative RK formulations and found that, in an oceanic framework, the gains over the scheme implemented here do not justify the added complexity. We develop the implicit formulation first followed by the split-explicit one.

3.4. Implicit free surface discretization

The implicit free surface approach is a suitable choice if stability is a primary concern, as the barotropic component has the desired property of being unconditionally stable. Campin et al. (2004) describe how to extend the implicit approach of Dukowicz and Smith (1994) to a nonlinear free surface. We illustrate the steps using a Forward-Euler discretization for illustrative purposes. In an implicit free-surface discretization we have,

$$\frac{\eta^{n+1} - \eta^n}{\Delta t} = -\boldsymbol{\nabla} \cdot \int_{-H}^{0} \sigma^{n+1} \mathbf{u}^{n+1} \, \mathrm{d}r \,, \tag{17}$$

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\Delta t} = \mathbf{G}_u^n - g \, \boldsymbol{\nabla} \eta^{n+1} \,. \tag{18}$$

Substituting $\mathbf{u}^{n+1}$ from (18) into (17), we get,

$$\frac{\eta^{n+1} - \eta^n}{\Delta t} = -\boldsymbol{\nabla} \cdot \int_{-H}^{0} \sigma^{n+1} \big(\underbrace{\mathbf{u}^n + \Delta t \, \mathbf{G}_u^n}_{\overset{\text{def}}{=} \mathbf{u}^*} \big) \, \mathrm{d}r + \Delta t \, g \, \boldsymbol{\nabla} \cdot \int_{-H}^{0} \sigma^{n+1} \boldsymbol{\nabla} \eta^{n+1} \, \mathrm{d}r \,, \tag{19}$$

where $\mathbf{u}^*$ is the predictor velocity at the next time step. After some reordering, and noticing that η does not depend on r,

$$\eta^{n+1} - \Delta t^2 g \, \boldsymbol{\nabla} \cdot \big(\mathcal{H}^{n+1} \boldsymbol{\nabla} \eta^{n+1} \big) = \eta^n - \Delta t \, \boldsymbol{\nabla} \cdot \underbrace{\int_{-H}^{0} \sigma^{n+1} \mathbf{u}^* \, \mathrm{d}r}_{\overset{\text{def}}{=} \mathbf{U}^*} \,. \tag{20}$$

Above, we denoted the vertically-integrated predictor velocity with $\mathbf{U}^*$. The grid needs to remain constant through the free surface update to be able to express (20) as a linear system. Therefore, the free surface consistency condition imposes a staggering between grid evolution and free surface evolution, i.e.,

$$\mathcal{H}^{n+1} = H + \eta^n \,, \tag{21}$$

This staggering is also crucial to ensure local tracer conservation, as the barotropic velocity, $\mathbf{U}^n = \int_{-H}^{0} \sigma^n \mathbf{u}^n \, \mathrm{d}r$, advects both the grid and the tracers. An RK method cannot use a staggering approach at the stage level, as it would intertwine different RK stages and would result in

$$\sigma^{n+1} = \frac{H + \eta^{n,M-1}}{H} \tag{22}$$

where $\eta^{n,M-1}$ is the free surface at the second-to-last RK stage at time step n. Since RK stages are not physical time steps, $\eta^{n,M-1}$ is an auxiliary variable, not suitable to use to estimate the physical grid stretching at timestep σ^{n+1}. There are two possibilities: (1) leaving the grid frozen during the RK stages and updating σ^{n+1} at the last stage or (2) ensuring consistent evolution of grid thickness and free surface during the stages and correcting the tracer transport to satisfy local tracer conservation. We take approach (2), for which, in all stages

$$\sigma^{m+1} = \frac{H + \eta^{m+1}}{H}\ , \quad \text{for} \ \ m = 0, \dots, M-1\ . \tag{23}$$

For simpler notation, we denote hereafter the RK substep size

$$\gamma_m \overset{\text{def}}{=} \frac{\Delta t}{M - m}\ , \quad \text{for} \ \ m = 0, \dots, M-1\ . \tag{24}$$

To satisfy local tracer conservation, we advect the tracers with corrected transport velocities $\mathbf{u}^\dagger$ and $\omega^\dagger$ that reproduce the barotropic transport that evolves the free surface from η^n to η^{m+1} over the substep size γ_m, which, for the implicit discretization, satisfies

$$\frac{\eta^{m+1} - \eta^n}{\gamma_m} = -\boldsymbol{\nabla} \cdot \mathbf{U}^\dagger\ , \tag{25}$$

where we define

$$\mathbf{U}^\dagger \overset{\text{def}}{=} \int_{-H}^{0} \sigma^m \mathbf{u}^{m+1} \mathrm{d}r\ , \tag{26}$$

such that $\mathbf{u}^\dagger$ is the velocity at stage m corrected by the barotropic transport at stage $m+1$. If applied to the RK method described in the previous section, a substep of the coupled RK - implicit free surface (RK-IM) is formulated as a predictor-corrector algorithm that follows

$$\mathbf{u}^* = \mathbf{u}^n + \gamma_m\, \mathbf{G}_u^m\ , \tag{27}$$

$$\eta^{m+1} - \gamma_m^2 g\, \boldsymbol{\nabla} \cdot \left(\mathcal{H}^m \boldsymbol{\nabla} \eta^{m+1}\right) = \eta^n - \gamma_m \boldsymbol{\nabla} \cdot \mathbf{U}^*\ , \tag{28}$$

$$\mathbf{u}^{m+1} = \mathbf{u}^* - \gamma_m g\, \boldsymbol{\nabla} \eta^{m+1}\ , \tag{29}$$

$$\sigma^{m+1} = \frac{H + \eta^{m+1}}{H}\ , \tag{30}$$

$$\mathbf{u}^\dagger = \mathbf{u}^m - \frac{1}{\mathcal{H}^m} \int_{-H}^{0} \sigma^m (\mathbf{u}^m - \mathbf{u}^{m+1})\, \mathrm{d}r \tag{31}$$

$$\omega^\dagger = -\int_{-H}^{r} \left[\boldsymbol{\nabla} \cdot (\sigma^m \mathbf{u}^\dagger) + \frac{1}{H} \boldsymbol{\nabla} \cdot \mathbf{U}^\dagger \right] \mathrm{d}r'\ , \tag{32}$$

$$C^{m+1} = \frac{\sigma^n}{\sigma^{m+1}} C^n - \frac{\gamma_m}{\sigma^{m+1}} \left[\boldsymbol{\nabla} \cdot (\sigma^m \mathbf{u}^\dagger C^m) + \frac{\partial(\omega^\dagger C^m)}{\partial r} \right]\ , \tag{33}$$

$$\omega^{m+1} = -\int_{-H}^{r} \left[\boldsymbol{\nabla} \cdot (\sigma^{m+1} \mathbf{u}^{m+1}) + \frac{1}{H} \boldsymbol{\nabla} \cdot \mathbf{U}^{m+1} \right] \mathrm{d}r'\ . \tag{34}$$

The algorithm is summarized in 1.

3.5. Split–explicit free surface discretization

The split explicit free surface discretization is a method that explicitly solves for the fast (barotropic) timescales within a large (baroclinic) timestep through a sub-cycling procedure. The sub-cycling here is distinct from the RK stages. Using a Forward Euler for the baroclinic mode and solving the barotropic mode using a

Algorithm 1 Flowchart of the substep of the Runge–Kutta(RK) algorithm with implicit free surface (RK-IM). All variables carry both the time-step index n and the RK stage index m, with $\varphi^{n,0} \equiv \varphi^n$ and $\varphi^{n,M} \equiv \varphi^{n+1}$. A superscript asterisk marks a predictor quantity; a superscript dagger marks a quantity corrected for tracer conservation.

```
Input: $\mathbf{u}^{n,0}$, $\mathbf{u}^{n,m}$, $\eta^{n,0}$, $\sigma^{n,0}$, $\sigma^{n,m}$, $C^{n,0}$, $C^{n,m}$, $\mathbf{G}_u^{n,m}$, $\gamma_m$
Predict velocity: $\mathbf{u}^* = \mathbf{u}^{n,0} + \gamma_m \mathbf{G}_u^{n,m}$
Solve for $\eta^{n,m+1}$: $\eta^{n,m+1} - \gamma_m^2 g\, \nabla \cdot (\mathcal{H}^{n,m} \nabla \eta^{n,m+1}) = \eta^{n,0} - \gamma_m \nabla \cdot \mathbf{U}^*$
Correct velocity: $\mathbf{u}^{n,m+1} = \mathbf{u}^* - \gamma_m g\, \nabla \eta^{n,m+1}$
Update thickness: $\sigma^{n,m+1} = (H + \eta^{n,m+1})/H$
Compute transport velocity: $\mathbf{u}^\dagger = \mathbf{u}^{n,m} - \left[\int_{-H}^{0} \sigma^{n,m} (\mathbf{u}^{n,m} - \mathbf{u}^{n,m+1})\, \mathrm{d}r\right] / \mathcal{H}^{n,m}$
Diagnose $\omega^\dagger$ from continuity
Advect tracers: $C^{n,m+1} = (\sigma^{n,0}/\sigma^{n,m+1}) C^{n,0} - (\gamma_m/\sigma^{n,m+1})\, \nabla \cdot (\sigma^{n,m} \mathbf{u}^\dagger C^{n,m})$
Diagnose $\omega^{n,m+1}$ from continuity
Output: $\mathbf{u}^{n,m+1}$, $\eta^{n,m+1}$, $\sigma^{n,m+1}$, $C^{n,m+1}$, $\omega^{n,m+1}$
```

forward-backward scheme, the free surface evolution follows

$$\eta^{n,0},\ \mathbf{U}^{n,0} = \eta^n, \tag{35}$$

$$\frac{\mathbf{U}^{n,s+1} - \mathbf{U}^{n,s}}{\gamma_B} = \int_{-H}^{0} \sigma^n \mathbf{G}_u^n \, \mathrm{d}r - g \mathcal{H}^{n,s} \nabla \eta^{n,s} \quad \text{for } s = 0, \dots, N_S\,, \tag{36}$$

$$\frac{\eta^{n,s+1} - \eta^{n,s}}{\gamma_B} = -\nabla \cdot \mathbf{U}^{n,s+1} \quad \text{for } s = 0, \dots, N_S\,, \tag{37}$$

where s indicates the sub-cycling index, $\gamma_B \overset{\text{def}}{=} 2\Delta t / N_S$ is the barotropic sub-cycling time step size, N_S the total number of sub-cycles, and $\sigma^n \mathbf{G}_u^n$ is computed at the current baroclinic step (or substep in the case of the RK time discretization) and frozen during sub-cycling. To ensure the stability of this scheme, we use the filtering approach of Shchepetkin and McWilliams (2005), where the updated barotropic variables are computed from

$$\eta^{n+1} = \sum_{s=1}^{\tilde{N}_S} \alpha^{n,s} \eta^{n,s}\,, \quad \mathbf{U}^{n+1} = \sum_{s=1}^{\tilde{N}_S} \alpha^{n,s} \mathbf{U}^{n,s}\,. \tag{38}$$

Above, α are filtering weights, and $\tilde{N}_S \geq N_S/2$ is the final sub-cycling index corresponding to the first index for which $\alpha^s \leq 0$. The α weights are centered on $n+1$, and integrate to one. The total sub-cycles $\tilde{N}_S$ depends on the particular choice of filtering weights; therefore, in the remainder of the paper, we indicate sums up to N_S for notational convenience by implying $\alpha^s = 0$ for $s > \tilde{N}_S$. Hereafter we omit the n index in subcycled variables, assuming we are always referring to a time step n. In terms of free surface evolution, the whole sub-cycling procedure can be summarized by

$$\frac{\eta^{n+1} - \eta^n}{\Delta t} = G_\eta\,, \tag{39}$$

where the tendency G_η can be again expressed, as we have done in (25), as the divergence of a corrected barotropic transport $\mathbf{U}^\dagger$,

$$G_\eta = -\nabla \cdot \mathbf{U}^\dagger\,. \tag{40}$$

Following Shchepetkin and McWilliams (2005), we can compute the corrected transport for a split explicit free surface discretization as

$$\mathbf{U}^\dagger = \sum_{s=1}^{N_S} \beta^s \mathbf{U}^s\,, \tag{41}$$

where the weights β are the barotropic transport weights, complementary to the α weights, centered on $n + 1/2$. These two sets of weights are related via:

$$\beta^s = \frac{2}{N_S} \sum_{s'=s}^{N_S} \alpha^{s'}\,. \tag{42}$$

Algorithm 2 Flowchart of the substep of the Runge–Kutta(RK) algorithm with split–explicit free surface (RK-SE). All variables carry both the time-step index n and the RK stage index m, with $\varphi^{n,0} \equiv \varphi^{n}$ and $\varphi^{n,M} \equiv \varphi^{n+1}$. A superscript asterisk marks a predictor quantity; a superscript dagger marks a quantity corrected for tracer conservation.

```
Input: $\mathbf{u}^{n,0}$, $\mathbf{u}^{n,m}$, $\eta^{n,0}$, $\mathbf{U}^{n,0}$, $\sigma^{n,0}$, $\sigma^{n,m}$, $C^{n,0}$, $C^{n,m}$, $\mathbf{G}_u^{n,m}$, $\gamma_m$
Barotropic sub-cycling: $\eta^{n,m+1}, \mathbf{U}^{n,m+1}, \mathbf{U}^{\dagger} = \mathcal{B}(\eta^{n,0}, \mathbf{U}^{n,0}, \mathbf{G}_u^{n,m}, \gamma_m)$
Compute transport velocity: $\mathbf{u}^{\dagger} = \mathbf{u}^{n,m} - \left(\int_{-H}^{0} \sigma^{n,m}\mathbf{u}^{n,m}\,\mathrm{d}r - \mathbf{U}^{\dagger}\right)/\mathcal{H}^{n,m}$
Diagnose $\omega^{\dagger}$ from continuity
Update thickness: $\sigma^{n,m+1} = (H + \eta^{n,m+1})/H$
Advect tracers: $C^{n,m+1} = (\sigma^{n,0}/\sigma^{n,m+1})C^{n,0} - (\gamma_m/\sigma^{n,m+1})\,\nabla\cdot(\sigma^{n,m}\mathbf{u}^{\dagger}C^{n,m})$
Predict velocity: $\mathbf{u}^{*} = \mathbf{u}^{n,0} + \gamma_m \mathbf{G}_u^{n,m}$
Reconcile with barotropic: $\mathbf{u}^{n,m+1} = \mathbf{u}^{*} - \left(\int_{-H}^{0} \sigma^{n,m+1}\mathbf{u}^{*}\,\mathrm{d}r - \mathbf{U}^{n,m+1}\right)/\mathcal{H}^{n,m+1}$
Diagnose $\omega^{n,m+1}$ from continuity
Output: $\mathbf{u}^{n,m+1}$, $\eta^{n,m+1}$, $\mathbf{U}^{n,m+1}$, $\sigma^{n,m+1}$, $C^{n,m+1}$, $\omega^{n,m+1}$
```

Once again, given equation (39), to ensure local tracer conservation we must advect tracers using a corrected transport velocity $\mathbf{u}^{\dagger}$ that satisfies

$$\int_{-H}^{0} \sigma^{n}\mathbf{u}^{\dagger} dr = \mathbf{U}^{\dagger} \tag{43}$$

such that the tracer evolution for a constant tracer matches the discretized continuity equation. We can summarize the barotropic sub-cycling step as a function $\mathcal{B}$ that, given the barotropic state at time step n and the "slow" tendency term, returns the barotropic state at $n+1$ and the averaged barotropic transport that evolves the free surface:

$$\eta^{n+1}, \mathbf{U}^{n+1}, \mathbf{U}^{\dagger} = \mathcal{B}\left(\eta^{n}, \mathbf{U}^{n}, \mathbf{G}_u^{n}, \Delta t\right) \ . \tag{44}$$

Moving to the RK discretization for the baroclinic mode, we can design a scheme that satisfies all three desired numerical properties by performing a barotropic step at each RK stage. We apply one barotropic substep, complete with barotropic–baroclinic reconciliation, at the substep level. This ensures the desired consistency between the tendency $\mathbf{G}_u^{m}$ used in the baroclinic velocity evolution and the "slow" tendency seen by the barotropic velocity, which would be lost if the coupling were not done at the substep level. Therefore, if applied to the RK method described in section 2, a substep of a paired RK–split explicit free surface (RK-SE), reads

$$\eta^{m+1}, \mathbf{U}^{m+1}, \mathbf{U}^{\dagger} = \mathcal{B}\left(\eta^{n}, \mathbf{U}^{n}, \mathbf{G}_u^{m}, \gamma_m\right) \ , \tag{45}$$

$$\mathbf{u}^{\dagger} = \mathbf{u}^{m} - \frac{1}{\mathcal{H}^{m}}\left(\int_{-H}^{0} \sigma^{m}\mathbf{u}^{m}\,\mathrm{d}r - \mathbf{U}^{\dagger}\right) , \tag{46}$$

$$\omega^{\dagger} = -\int_{-H}^{r}\left[\nabla\cdot(\sigma^{m}\mathbf{u}^{\dagger}) + \frac{1}{H}\nabla\cdot\mathbf{U}^{\dagger}\right]\mathrm{d}r' \ , \tag{47}$$

$$\sigma^{m+1} = \frac{H + \eta^{m+1}}{H} \ , \tag{48}$$

$$C^{m+1} = \frac{\sigma^{n}}{\sigma^{m+1}}C^{n} - \frac{\gamma_m}{\sigma^{m+1}}\left[\nabla\cdot(\sigma^{m}\mathbf{u}^{\dagger}C^{m}) + \frac{\partial(\omega^{\dagger}C^{m})}{\partial r}\right] , \tag{49}$$

$$\mathbf{u}^{*} = \mathbf{u}^{n} + \gamma_m\mathbf{G}_u^{m} \ , \tag{50}$$

$$\mathbf{u}^{m+1} = \mathbf{u}^{*} - \frac{1}{\mathcal{H}^{m+1}}\left(\int_{-H}^{0} \sigma^{m+1}\mathbf{u}^{*}\,\mathrm{d}r - \mathbf{U}^{m+1}\right) , \tag{51}$$

$$\omega^{m+1} = -\int_{-H}^{r}\left[\nabla\cdot(\sigma^{m+1}\mathbf{u}^{m+1}) + \frac{1}{H}\nabla\cdot\mathbf{U}^{m+1}\right]\mathrm{d}r' \ . \tag{52}$$

This time discretization is designed so that the tracer is both locally and globally conserved, while the consistency between the thickness and free surface evolutions is ensured by (48). These steps are summarized in algorithm 2.

4. Linear stability analysis

Next we must demonstrate the robustness of the schemes designed in section 3. To do so, we compute the linear stability of both RK-IM and RK-SE schemes. To simplify the analysis, we assume two-dimensional dynamics in the x–z plane and hydrostatic balance. Under linearization, we lose self-advection of momentum, and the advection of buoyancy is limited to wN^2, where $N^2 \overset{\text{def}}{=} \partial_z \overline{b}$ is a constant background stratification. We also linearize the free surface equations around the rest state, which implies that the grid does not follow the free surface ($\sigma = 1$). The derivation follows Demange et al. (2019) and project variables of into vertical modes ϕ_q with details provided in Appendix B. Here, we present only the key discretized equations and stability results. Within this section and Appendix B we are tracking the evolution within a single time step, so we drop the leading time-step index n and write φ^m in place of $\varphi^{n,m}$. We retain the asterisk and dagger conventions introduced in section 3, with a superscript asterisk denoting a predictor quantity (u^*) and a superscript dagger denoting a transport-corrected quantity ($u^\dagger$, $w^\dagger$). For cleaner notation, we also omit the stage subscript from the RK substep size γ_m defined in (24) and simply write γ. The linearized RK-IM substep reads

$$u^* = u^n - \gamma \frac{\partial p^m}{\partial x} \,, \tag{53}$$

$$b^{m+1} = b^n - \gamma N^2 w^p \,, \tag{54}$$

$$\eta^{m+1} - \gamma^2 g H \frac{\partial^2 \eta^{m+1}}{\partial x^2} = \eta^n - \gamma \int_{-H}^{0} \frac{\partial u^*}{\partial x} \, \mathrm{d}r \,, \tag{55}$$

$$u^{m+1} = u^* - \gamma g \frac{\partial \eta^{m+1}}{\partial x} \,. \tag{56}$$

While the linearized RK-SE scheme yields

$$U^{s+1} = U^s - \gamma_B \left(gH \frac{\partial \eta^s}{\partial x} + \int_{-H}^{0} \frac{\partial p^m}{\partial x} \, \mathrm{d}z \right) , \quad \text{with} \quad U^{s=0} = U^n \,, \tag{57}$$

$$\eta^{s+1} = \eta^s - \gamma_B \frac{\partial U^{s+1}}{\partial x} \,, \quad \text{with} \quad \eta^{s=0} = \eta^n \,, \tag{58}$$

$$U^{m+1} = \sum_{s=1}^{N_S} \alpha^s U^s \,, \tag{59}$$

$$\eta^{m+1} = \sum_{s=1}^{N_S} \alpha^s \eta^s \,, \tag{60}$$

$$b^{m+1} = b^n - \gamma N^2 w^p \,, \tag{61}$$

$$u^* = u^n - \gamma \frac{\partial p^m}{\partial x} \,, \tag{62}$$

$$u^{m+1} = u^* - \frac{1}{H} \left(\int_{-H}^{0} u^* \, \mathrm{d}z + U^{m+1} \right) . \tag{63}$$

In the buoyancy update equations (54) and (61), w^p is an estimate of the vertical velocity that advects the tracers. Using the velocity from the previous step in an RK framework would yield $w^p = w^m$. However, we use the vertical velocity diagnosed by the corrected transport velocities ($\mathbf{u}^\dagger$) computed in equations (25) and (46) for the implicit and split–explicit discretization, respectively. In this linearized framework,

$$w^p = - \int_{-H}^{z} \frac{\partial u^\dagger}{\partial x} \, \mathrm{d}z \,, \tag{64}$$

where

$$\int_{-H}^{0} u^\dagger \, \mathrm{d}z = \int_{-H}^{0} u^{m+1} \, \mathrm{d}z \quad \text{for an implicit free surface,} \tag{65}$$

$$\int_{-H}^{0} u^\dagger \, \mathrm{d}z = \sum_{s=1}^{N_S} \beta^s U^s \qquad \text{for a split–explicit free surface.} \tag{66}$$

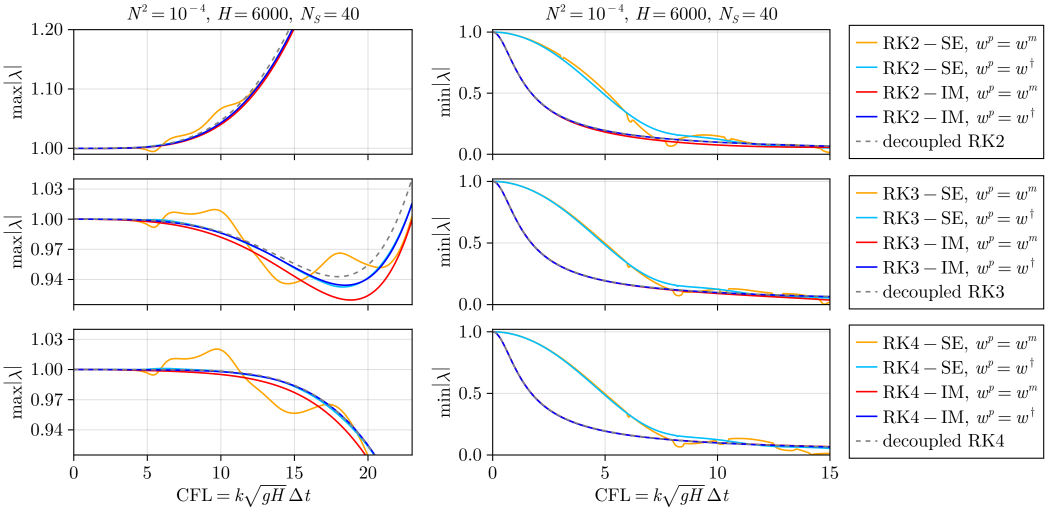

Figure 2: Maximum and minimum eigenvalues of the different evolution matrices computed in Appendix B for RK two to four stages. The baroclinic mode typically corresponds to the eigenvalue with the largest magnitude, while the barotropic mode corresponds to the smallest eigenvalue.

As explained in section 3, we have chosen to use this vertical velocity for tracer conservation reasons; however, as we demonstrate in this section, this choice has other benefits: it decreases the baroclinic dissipation of the implicit free surface discretization and enables a baroclinically stable RK substepping procedure for a split–explicit discretization, which would not be possible using $w^p = w^m$. The modal decomposition and matrix formulation for the systems shown above (including both $w^p = w^m$ and $w^p = w^\dagger$) are detailed in Appendix B. Figure 2 shows the magnitude of the largest and smallest eigenvalues of the four different evolution matrices derived in Appendix B, for a two-stage, a three-stage, and a four-stage RK scheme. The analysis is performed in typical ocean conditions, with a constant stratification of $N^2 = 10^{-4}$ s^{-2} and a bottom height of $H = 6000$ m. For the split–explicit discretization, we use $N_S = 40$ substeps and the low-dispersion filter of Shchepetkin and McWilliams (2005), defined by

$$\alpha(\tau) = \left(\frac{\tau}{\tau_0}\right)^p \left[1 - \left(\frac{\tau}{\tau_0}\right)^q\right] - r\left(\frac{\tau}{\tau_0}\right)\,, \quad \text{with} \quad \begin{cases} \tau_0 = \dfrac{(p+2)(p+q+2)}{(p+1)(p+q+1)}\,, \\ p = 2\,, \\ q = 4\,, \\ r = 0.18927\,. \end{cases} \tag{67}$$

Within this framework, it is easy to change the filter or use a different barotropic time discretization, such as the dissipative forward-backward scheme described by Demange et al. (2019). We implemented various filters and the dissipative forward-backward scheme in Oceananigans. However, we found the dispersion from the dissipative forward-backward scheme too large to produce acceptable solutions, even in simplified test cases. The results of the stability analysis for the baroclinic mode differ only slightly from the amplification factors of an uncoupled RK scheme. RK2 is unconditionally unstable, confirming that the dissipation introduced in the barotropic solver is insufficient to stabilize the baroclinic mode. RK3 and RK4 follow curves qualitatively similar to the ones in figure 1, with slightly lower λ damped by the barotropic coupling. However, figure 2 demonstrates the importance of selecting the appropriate transport velocity when coupling the barotropic and baroclinic modes. In the RK-IM case, both w^m and $w^\dagger$ lead to a stable discretization with very little impact on the barotropic mode's amplification factor, but using $w^\dagger$ improves the baroclinic amplification, reducing the impact of the implicit dissipation on this mode. In the RK-SE case, $w^p = w^\dagger$ yields a clean discretization with baroclinic mode amplification similar to that obtained in section 2. In contrast, $w^p = w^m$ limits the overall scheme's robustness: baroclinic instabilities grow with increasing stratification and domain depth, indicating that $w^p = w^m$ is an incorrect coupling choice for multi-stage schemes like RK.

case	time discretization	tracer advection	free surface
RK3-SE	RK3	WENO, 7th order	split explicit
RK3-IM	RK3	WENO, 7th order	implicit
AB2-SE	AB2	WENO, 7th order	split explicit
RK3-UP	RK3	Upwind, 3rd order	split explicit

Table 1: Numerical details of the different test cases employed.

To summarize, with $w^p = w^\dagger$ (the only configuration considered in the following results), RK-IM is much more dissipative than RK-SE for barotropic motions, while the two schemes produce comparable damping of the baroclinic mode across the stable CFL range.

5. Results and validation

We further test the RK-IM and RK-SE formulations with a suite of idealized test cases, designed to highlight differences in the discretization's performance across different dynamical regimes. We assess the numerical discretization's proficiency based on the resolved kinetic energy and the numerical mixing introduced. To estimate the numerical mixing, we implement the method described by Banerjee et al. (2024) that provides an exact value (local in time and space) for the numerically induced dissipation of buoyancy variance in the x, y, and z direction (P_x, P_y, and P_z). Since this methodology was initially devised for single-step, single-stage time discretizations, we extend it to the AB2 scheme currently implemented in Oceananigans, and to the RK scheme developed in this paper. A detailed description is provided in Appendix C. With the numerical variance dissipation, we estimate the numerical diffusivity as

$$\kappa_{\text{num}} = -\frac{1}{2}\left(\overline{P}_x + \overline{P}_y + \overline{P}_z\right)\left(\overline{|\nabla b|^2}\right)^{-1} , \tag{68}$$

where the bar indicates an averaging operator (either time or space averaging, depending on the test case). Averaging is introduced to avoid divisions by zero when the flow is locally unstratified. The details of the tested numerical discretizations are shown in Table 1. We use a 7th-order WENO reconstruction for tracer advection and, where not explicitly stated, the momentum advection uses the WENO vector-invariant 9th-order scheme, as described by Silvestri et al. (2024). We include an additional case, RK3-UP, with a lower-order tracer advection scheme to help separate and distinguish the numerical mixing caused by time and space discretization.

5.1. Internal tide

The internal tide configuration mirrors the Oceananigans documentation example of the same name (Climate Modeling Alliance & contributors, 2026). It consists of a two-dimensional domain, 2-kilometer deep, initialized with a constant stratification $N^2 = 10^{-4}$ s^{-2}, and forced by a zonal oscillatory tide applied as a body forcing:

$$F_U = \frac{\omega^2 - f^2}{\omega} U \sin(\omega t) , \quad \text{with} \quad \begin{cases} U &= \epsilon \omega D , \\ \epsilon &= 0.1 , \\ D &= 20 \text{ km} , \\ \omega &= 2\pi/T , \\ T &= 12.421 \text{ hr} . \end{cases} \tag{69}$$

The forced tide flows over a 250 meter-tall Gaussian seamount represented with an immersed-boundary method. The horizontal and vertical resolutions are 7.8125 km and 16.625 m (128 layers spanning 2 km), respectively. This test case isolates the role of time discretization: the spatial advection operator, which is usually dominant in causing numerical mixing, plays a secondary role because this setup is mostly linear. Therefore, we expect to see results that closely match the analytical predictions derived in section 4. We use a time step size of 5 minutes for AB2 and 10 minutes for the RK3 schemes, with $N_S = 60$ for the split–explicit

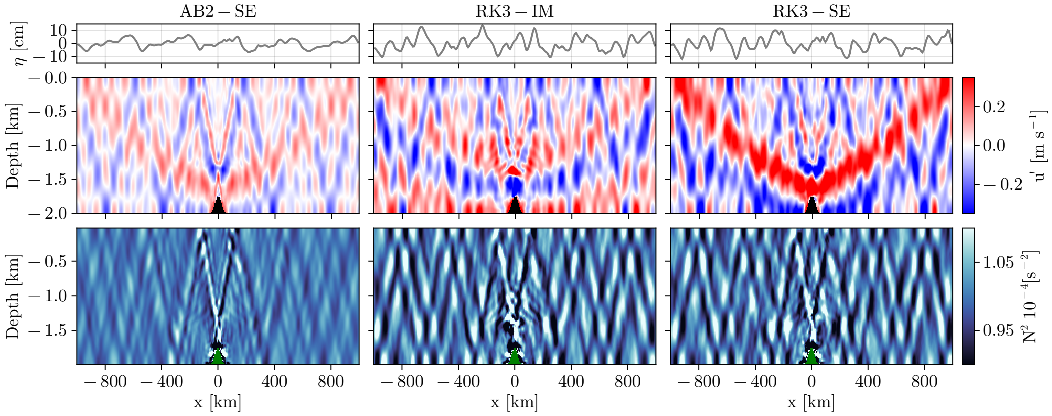

Figure 3: The internal tide configuration after 40 days of evolution as computed by the different time discretization schemes. The top panels show the free surface elevation, while the center and bottom panels show the instantaneous velocity fluctuations $u' = u - \int_V u \, \mathrm{d}V$ and the stratification $N^2 = \partial_z b$.

cases (independent of the global time step size). Simulations are initialized with $u = U$ and run for 40 days. No drag or explicit mixing is applied, so any tracer or momentum diffusion stems solely from the numerical mixing introduced by the time and space discretizations.

Figure 3 shows snapshots of free-surface elevation η, velocity fluctuations u', and stratification N^2 after 40 days. The RK3 solutions are more energetic than the AB2-SE one: both u' and N^2 exhibit sharper structures than in AB2-SE, indicating the lower implicit diffusivity introduced by RK3. However, the RK3-IM implicit free-surface treatment damps the solution relative to RK3-SE as demonstrated analytically in section 4. The free-surface elevation and stratification patterns in RK3-IM and RK3-SE are comparable; however, the associated velocity fields show phase shifts in the tidal response. Despite the differences in magnitude, the velocity fields are qualitatively very similar for AB2-SE and RK3-SE.

Figure 4 displays vertical profiles of the time-varying numerical diffusivity for all configurations. The diffusivity oscillates with the tidal cycle, peaking when the barotropic velocity is largest, yet AB2-SE consistently exhibits the highest values. Notably, AB2-SE remains more diffusive than the lower-order RK3-UP case, underscoring how, in this relatively linear test case, a dissipative time discretization alone can match or even surpass the mixing induced by a low-order, diffusive spatial discretization (Marchesiello et al., 2009; Schifano et al., 2025).

Figure 5 compares the volume-averaged turbulent kinetic energy, reference potential energy (RPE), available potential energy (APE), and free-surface variance, computed as

$$\mathrm{TKE} = \frac{1}{2V}\int_V \left({u'}^2 + w^2\right) \mathrm{d}V \,, \quad \mathrm{RPE} = \frac{1}{V}\int_V \rho z_{\mathrm{res}} \, \mathrm{d}V \,, \quad \mathrm{APE} = \frac{1}{V}\int_V \rho (z - z_{\mathrm{res}}) \, \mathrm{d}V \,, \tag{70}$$

where ρ is the in-situ density, $u' \overset{\mathrm{def}}{=} u - \int_V u \, \mathrm{d}V$, and z_{res} is the re-sorted height computed as in the work of Ilicak (2016) and Hill et al. (2012) (see figure 3 by Hill et al. (2012) for a visual description of z_{res}). The RPE metric tracks integrated mixing in the domain; since RPE can increase only due to numerical mixing, the higher the RPE, the greater the spurious diffusion introduced by the numerical discretization. RK3-SE and RK3-UP track each other closely in TKE, APE, and η^2, confirming that the numerical mixing introduced by the spatial discretization is secondary in this linear regime. RK3-IM is slightly more dissipative than RK3-SE because the implicit free-surface solve introduces stronger barotropic damping, which dominates in this linear regime. At the same time, AB2-SE shows the strongest decay in energetic quantities, again reflecting its higher numerical diffusivity. RPE behaves differently: despite its low TKE, AB2-SE shows the most significant RPE increase during the first 30 days, indicating that its diffusivity drives substantial irreversible mixing. After day 30, the RK3-UP case overtakes AB2-SE in RPE as the solution becomes more nonlinear and sensitive to the lower-order tracer advection. Notably, RK3-SE shows a smaller RPE increase

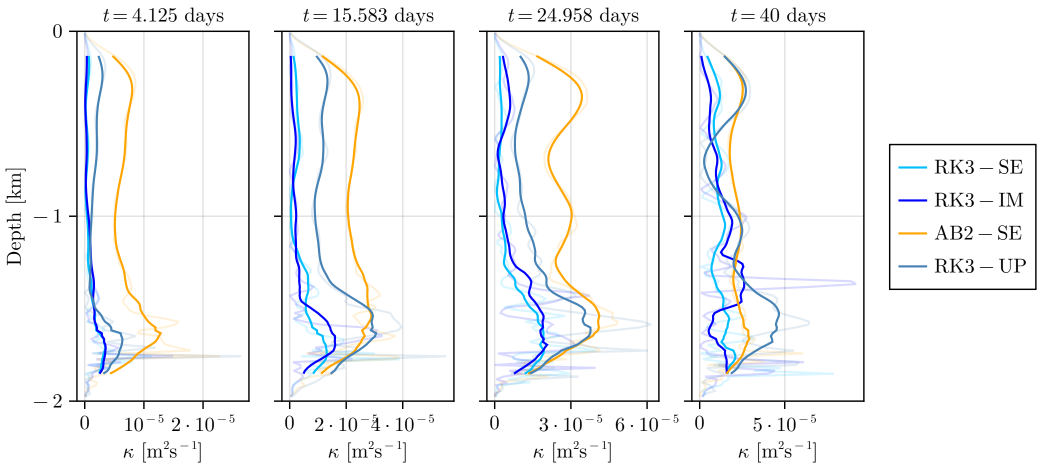

Figure 4: Total numerical diffusivity as a function of depth and time. The bar in this configuration indicates an average over the zonal x-direction. The transparent lines show the pointwise diffusivity computed via (68), while, to improve visualization, the solid lines show a running average of the same diffusivity over 8 vertical cells.

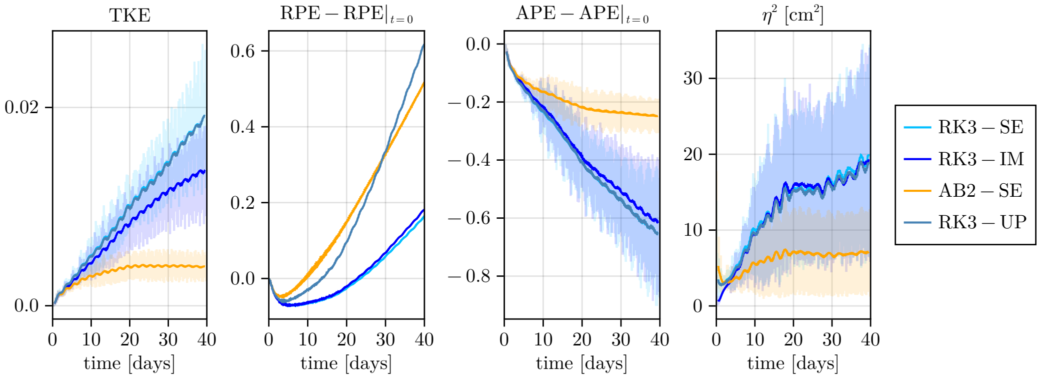

Figure 5: Energetics of the internal tide configuration for the different numerical discretizations. All quantities are volume averaged. The transparent lines in panels 1, 2, and 4 show the instantaneous values, while the solid lines show a 1-day running average.

than RK3-IM, despite the larger TKE and higher velocities, confirming that the numerical mixing in this setup is produced mainly by the time discretization.

5.2. Coastal baroclinic adjustment

We next examine a more complex yet still idealized problem: a coastal baroclinic adjustment configuration that resembles those described by Hetland (2017); Schlichting, Hetland, and Jones (2024). The test case is a rectangular domain of size 192 by 192 kilometers and 103 meters deep, with a uniform resolution of 2 kilometers in the horizontal and 1.7 meters in the vertical direction. We add a linearly sloping bathymetry, 5 meters deep on the southern edge, which reaches the bottom of the domain at 96 kilometers in the y-direction. The buoyancy is a linear function of temperature and salinity, where the initial hydrography is prescribed by

$$\begin{cases} T &= T_0 + \dfrac{N^2}{\alpha g} z \\ \\ S &= S_0 - \dfrac{S^2}{\beta g} z - \dfrac{M^2}{\beta g}(y_0 - y) \end{cases}, \quad \text{with} \quad \begin{cases} T_0 &= 25\,^\circ\mathrm{C} \\ S_0 &= 35 \text{ psu} \\ N^2 &= 10^{-4}\ \mathrm{s}^{-2} \\ S^2 &= 10^{-8}\ \mathrm{s}^{-2} \\ M^2 &= 1.2 \times 10^{-6}\ \mathrm{s}^{-2} \\ \alpha &= 1.7 \times 10^{-4}\ ^\circ\mathrm{K}^{-1} \\ \beta &= 7.6 \times 10^{-4}\ \mathrm{psu}^{-1} \\ y_0 &= 50\ \mathrm{km} \\ g &= 9.80665\ \mathrm{m\,s^{-2}} \end{cases} \tag{71}$$

This configuration is an "Eady problem with bathymetry"; in contrast to the previous case, it exhibits nonlinear dynamics, allowing baroclinic instability to develop parallel to the sloping bathymetry. With these dynamics, it provides another angle for assessing the accuracy of numerical discretizations since, as concluded by Schlichting et al. (2024), in this particular configuration, numerical mixing suppresses baroclinic instabilities. Therefore, the higher the numerical mixing, the slower the initial buoyancy gradient relaxes towards equilibrium. We also apply a quadratic bottom drag with a drag coefficient of 3×10^{-3} and a vertical Laplacian dissipation to the momentum with a viscosity of $\nu = 3 \times 10^{-5}\ \mathrm{m^2\,s^{-1}}$. We do not include any other form of explicit dissipation; hence, the dissipation of tracers is provided solely by numerical mixing. We use a time step of 5 minutes for AB2 and 10 minutes for the RK schemes, while for AB2-SE and RK3-SE, we compute N_S to ensure a maximum barotropic CFL of 0.7. We evolve the case from rest to 40 days, long-enough to generate a significant baroclinic adjustment of the initial meridional salinity gradient.

Figure 6 provides a qualitative view of the surface RK3-SE solution at day 40. From left to right, the figure shows relative vorticity and buoyancy dissipation in the x, y, and z directions. While some positive pockets of dissipation can be identified in all directions, denoting buoyancy variance production, most of the contours show negative values, indicating that, globally, numerical errors dissipate rather than amplify buoyancy variance. Interestingly, in this case, horizontal dissipation dominates over the vertical dissipation, with regions of intense dissipation occurring at the edge of the eddies.

Figure 7 shows buoyancy contours and zonally averaged salinity differences relative to RK3-IM. The left panel displays buoyancy contours for AB2-SE (dashed) and RK3-IM (solid), overlaid on the filled salinity difference $\overline{S}$(AB2-SE) − $\overline{S}$(RK3-IM); the center panel shows the same for RK3-SE. To demonstrate that these differences indicate an inhibition of submesoscale activity, which can be caused only by numerical mixing, the rightmost panel shows the zonally averaged salinity difference for RK3-IM between 24 and 40 days of evolution. The salinity difference patterns in the left and center panels closely resemble that in the rightmost panel, indicating that both AB2-SE and RK3-SE are effectively "temporally lagging" behind RK3-IM, with AB2-SE showing the largest lag. This is confirmed by the buoyancy contours: RK3-IM is the most stratified, followed closely by RK3-SE, with AB2-SE clearly lagging behind.

Figure 8 shows the diagnosed diffusivities where the averaging operator used, here, is an averaging in the zonal and meridional direction. Unlike the previous test case, RK3-IM and RK3-SE show similar diffusivity values, both consistently larger than AB2-SE. This apparent discrepancy can be attributed to the more ad-

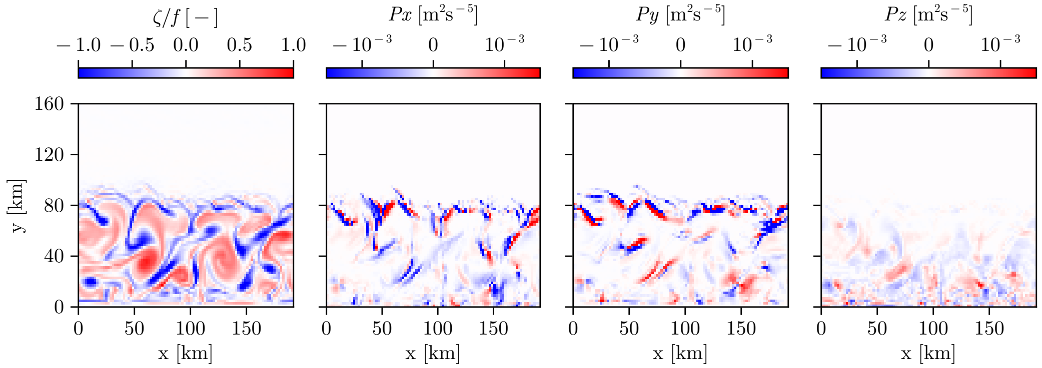

Figure 6: Instantaneous snapshots for surface variables after 40 days for case RK3-SE. From left to right: surface relative vorticity, and zonal, meridional and vertical numerical buoyancy dissipation, as defined in Appendix C in equations (181)–(183).

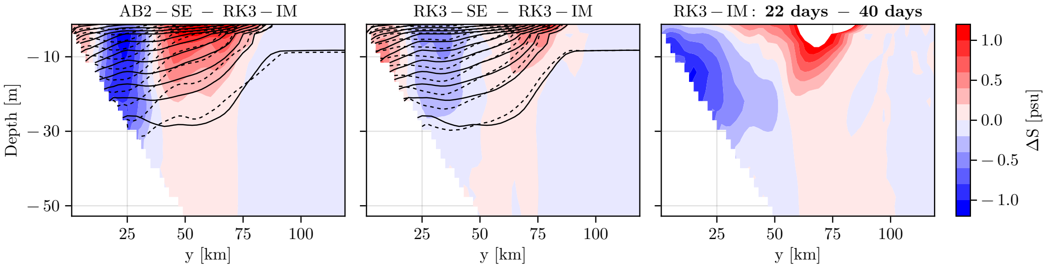

Figure 7: Zonally averaged buoyancy and salinity. The shaded contours show zonally averaged salinity differences; the contour lines show averaged buoyancy surfaces where the solid lines are RK3-IM. In the left panel, the salinity difference is computed between AB2-SE and RK3-IM; in the center panel, between RK3-SE and RK3-IM; and in the right panel, between RK3-IM evolved for 24 days and for 40 days.

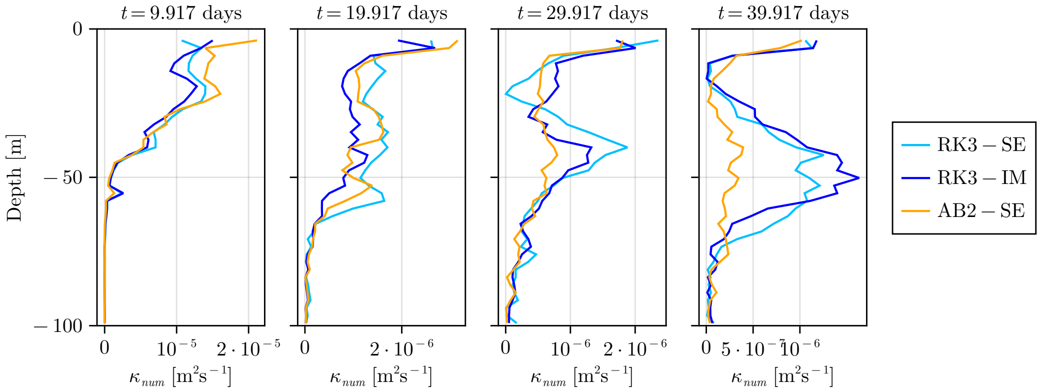

Figure 8: Numerical diffusivity computed via (68) as a function of depth and time. In (68), the bar denotes an average in time, and both zonal and meridional directions.

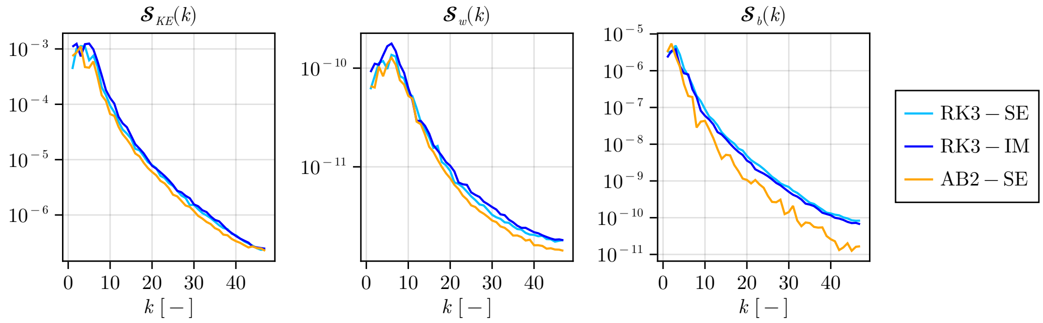

Figure 9: Power density spectra of kinetic energy (left), vertical velocity (center), and buoyancy (right), at 11 meters depth, time-averaged between days 24 and 40.

vanced state of the baroclinic instability in the RK cases, which generates stronger buoyancy gradients and, consequently, higher diagnosed diffusivities. The magnitude of κ_{num}, which is $O(10^{-6})$ or less, suggests that numerical buoyancy diffusivity does not play a dominant role in this case over the 40-day integration. This time scale is short enough that tracer diffusivity is likely less important than the effects of the implied numerical viscosity in the momentum field. This test case demonstrates that diagnosing numerical diffusivity alone may be insufficient to fully characterize the accuracy of a numerical discretization scheme.

Finally, Figure 9 shows zonal power density spectra averaged in the last 20 days, computed as follows:

$$\mathcal{S}_{KE}(k) = \frac{1}{2}\left(|\hat{u}(k)|^2 + |\hat{v}(k)|^2\right) , \quad \mathcal{S}_w(k) = |\hat{w}(k)|^2 , \quad \mathcal{S}_b(k) = |\hat{b}(k)|^2 , \tag{72}$$

where the hat indicates a Fourier transform. As expected, RK3-IM and RK3-SE retain appreciably more velocity variance in both the horizontal and vertical directions compared to AB2-SE, with RK3-IM showing slightly higher values across all spectra. AB2-SE shows signs of spectral blocking in the horizontal velocity spectra at small scales (the KE spectrum has an increased slope towards the largest wavenumbers), signaling the presence of possible grid-scale artifacts. In terms of buoyancy, AB2-SE displays the lowest variance across scales, once again highlighting the implicit diffusion introduced by the method.

5.3. Idealized channel flow

Our final test case is an idealized re-entrant channel with imposed mesoscale forcing and explicit subgrid mixing. Unlike previous transient configurations, this case relaxes toward statistical equilibrium, allowing us to assess how each time-stepping scheme affects the equilibrated solution and how numerical mixing interacts with a parameterized physical diffusivity. The configuration consists of a 1000-kilometer-long, 2000-kilometer-wide, 3-kilometer-deep re-entrant channel in the x direction and bounded in the y-direction. We use a 5-kilometer horizontal resolution and discretize the vertical direction with 90 non-uniformly spaced vertical layers. To account for vertical physics, we use the CATKE vertical mixing parameterization (Wagner, Hillier, et al., 2025) with a constant background diffusivity $\kappa = 10^{-5}$ m^2 s^{-1} and viscosity $\nu = 10^{-4}$ m^2 s^{-1}. The channel is forced by a sinusoidal zonal wind stress and a surface buoyancy flux,

$$\tau_x = -\tau_0 \sin\left(\frac{\pi y}{L_y}\right) , \quad Q_b = \begin{cases} Q_0 \cos\left(\dfrac{3\pi y}{L_y}\right) & y < \dfrac{5}{6} L_y , \\ 0 & \text{otherwise} , \end{cases} \tag{73}$$

with $\tau_0 = 0.1$ N m^{-2} and Q_0 corresponding to a maximum heat flux of 10 W m^{-2}. To obtain a non-trivial equilibrated solution, buoyancy is restored at the northern boundary within a sponge layer of width $L_y/10$, with a relaxation timescale of 7 days, toward the exponential profile

$$b_{\text{targ}}(z) = \Delta B \frac{e^{z/h} - e^{-L_z/h}}{1 - e^{-L_z/h}} , \tag{74}$$

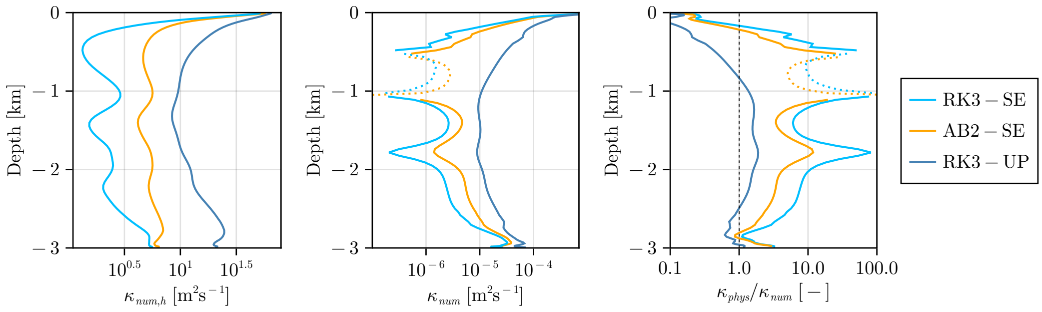

Figure 10: Numerical diffusivities (absolute values) computed via (68) for the equilibrated channel setup, where $\kappa_{\mathrm{num,h}}$ is the horizontal component of the diffusivity computed as $-(\overline{P}_x+\overline{P}_y)|\boldsymbol{\nabla}_h b|^2/2$. The bar denotes an average in time, and both zonal and meridional directions, while dotted lines show regions of negative numerical diffusivity (numerical variance production instead of destruction). In the right-most panel, the dashed vertical line marks $\kappa_{\mathrm{phys}}/\kappa_{\mathrm{num}} = 1$.

where $\Delta B = 8\alpha g$ and $h = 1000$ m. More details about this setup can be found in the papers by Abernathey, Marshall, and Ferreira (2011) and Hill et al. (2012). We use $\Delta t = 5$ min for the AB2 and $\Delta t = 10$ min for the RK3 time-discretization. To test the discretizations, we initialize the buoyancy with an equilibrated solution generated with the MITgcm (Marshall et al., 1997) in the same configuration and run for an additional 40 years. We then time-average the last 5 years (considered to be equilibrated to the new vertical mixing parameterization and numerical methods) and use it for our analyses below.

Figure 10 shows the diagnosed numerical diffusivity profiles for the equilibrated state, with panels from left to right showing the diffusivity in the horizontal direction ($x + y$), the total numerical diffusivity, and the ratio between the physical explicitly prescribed diffusivity and the total numerical diffusivity. While in the case above, diffusivity alone is insufficient, in this equilibrated case with explicit physical diffusivity, the comparison between physical and numerical diffusivity provides a meaningful metric of the performance of different configurations. For all the cases, the numerical diffusivity is $\sim$ 1–10 m^2 s^{-1} in the horizontal direction, while the total numerical diffusivity varies from 10^{-6} to 10^{-4} m^2 s^{-1}. In terms of absolute values, the largest numerical diffusivity is observed in the RK3-UP case, indicating the intense mixing that a low-order advection scheme can introduce. The AB2-SE scheme, which uses a 7th-order WENO for tracer advection, has lower numerical mixing than RK3-UP, but still larger than RK3-SE in the horizontal direction. While the physical mixing is roughly the same across all three cases (not shown), the ratio of physical to numerical mixing is different. For RK3-UP, $\kappa_{\mathrm{phys}}/\kappa_{\mathrm{num}} \lesssim 1$, indicating that the numerical discretization plays a large control over the effective mixing in the whole water column. AB2-SE and RK-SE show values much larger than one, reaching peaks > 10 in the interior of the domain, with RK-SE showing the largest values across the whole domain. These results highlight that in this case, where non-linearities dominate, the advection scheme plays a major role in the numerical diffusivity.

Figure 11 shows numerical dissipation averaged in the zonal direction split into horizontal and vertical dissipation. The contours in the figure are multiplied by $C = \exp{(-1.5z/1000)}$ to highlight deep regions of lower stratification. Positive values indicate regions of variance production, while negative values indicate variance destruction. For all discretizations, as in the previous case but unlike the internal tide test case, the horizontal direction dominates the numerical diffusivity. RK3-UP shows by far the largest variance dissipation, confirming the large diffusivity values presented in figure 10. AB2-SE also shows a much larger dissipation in the horizontal direction when compared to RK3-SE, explaining the much larger diffusivity in the left-most plot of figure 10. However, AB2-SE also shows a severe variance production in the vertical direction, which partly compensates for the variance destruction in the horizontal direction, and decreases the total numerical diffusivity to levels comparable to RK3-SE. On the contrary, RK3-SE has a much more benign behavior, with much lower variance destruction and production in the horizontal and vertical, respectively, when compared to AB2-SE.

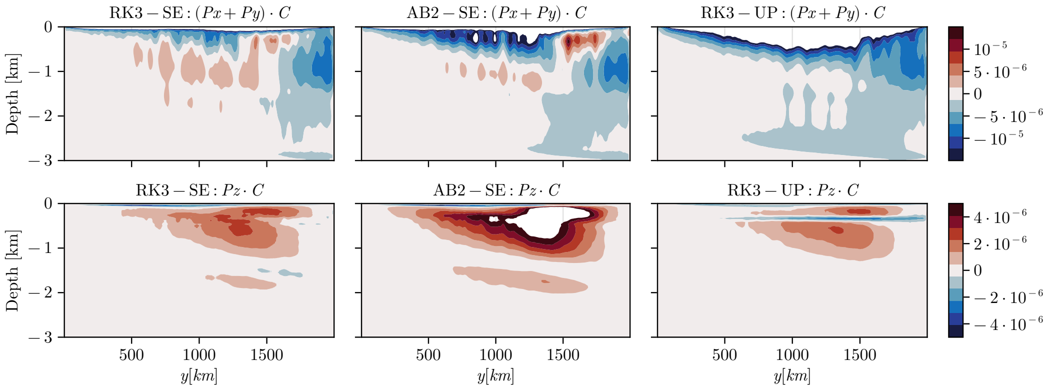

Figure 11: Numerical dissipation for the equilibrated channel setup, where the top row shows horizontal dissipation ($P_x + P_y$) while the bottom row shows the vertical dissipation. Both rows are multiplied by $C = \exp(-1.5z/1000)$ to highlight deep regions with lower stratification.

6. Conclusions

Baroclinic time-stepping in ocean models has long relied on low-order, single-stage methods. A common choice is the quasi Adams–Bashforth scheme, which, while computationally efficient, damps waves across typical CFL ranges, even when those waves are well resolved. This dissipation generates spurious mixing that degrades solution quality, especially over long extended integrations, yet the contribution from temporal discretization has remained largely unquantified.

As a lower-dissipation alternative, we developed here a low-storage Runge–Kutta framework that satisfies three essential constraints for free-surface ocean models: *(i)* consistency between free-surface evolution and grid geometry, *(ii)* exact tracer conservation, and *(iii)* robust stability across coupled barotropic–baroclinic modes. Both implicit and split–explicit free-surface treatments achieve these goals. Our stability analysis revealed that to maintain baroclinic stability, the split–explicit RK implementations require that the tracers be advected by the averaged barotropic transport velocity $w^\dagger$ rather than the velocity from the previous substep w^m.

To demonstrate the benefits of this time-discretization, we validated our developed frameworks through three idealized test cases. In a linear internal tide problem, temporal discretization effects rival those from low-order spatial schemes, confirming that dissipative time-stepping alone can dominate numerical mixing (Marchesiello et al., 2009; Schifano et al., 2025). In a nonlinear coastal baroclinic adjustment, reduced dissipation preserves physical instabilities that AB2 suppresses. RK3-IM maintains the sharpest stratification, followed closely by RK3-SE, both showing substantially stronger submesoscale variability than AB2 (Schlichting et al., 2024). In an equilibrated re-entrant channel, RK3 shows more beneficial numerical properties than AB2 with lower variance destruction and production throughout the water column, leading to lower numerical diffusivities and higher physical to numerical diffusivity ratios.

The above-mentioned results translate directly into practice. We implemented the RK discretization in Oceananigans as an alternative to the available Adams–Bashforth (Wagner, Silvestri, et al., 2025; Silvestri et al., 2025), offering users more accurate simulations with minimal additional cost. Because our RK formulation modifies only the baroclinic tendency evaluation, retrofitting existing split–explicit models requires minimal code changes. Additionally, the combination of modal stability analysis with exact tracer-variance budgets provides an evaluation methodology transferable to other ocean models. For climate-scale applications, maintaining stratification and preserving submesoscale variability becomes increasingly critical as models target higher resolutions (Hill et al., 2012).

Several extensions merit investigation: optimizing IMEX pairings to balance the trade-offs between implicit and split–explicit treatments, leveraging variance budgets for adaptive timestep control, and generalizing the framework to non-hydrostatic dynamics. Validation in realistic global configurations, where complex forcing, topography, and coupled biogeochemistry introduce additional challenges, will ultimately determine the full benefits of low-storage RK methods for operational climate simulations.

Acknowledgments

We thank Milan Klöwer for helpful discussions on the numerical precision of the split RK3 scheme and Siddhartha Bishnu for comments on an early draft of this paper. This project is supported by Schmidt Sciences under the Climate Modeling Alliance project. In addition, S.S. is supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101209201 and N.C.C. is supported by the Australian Research Council under the Center of Excellence for the Weather of the 21st Century CE230100012 and the Discovery Project DP240101274.

Open Research Section

The RK-IM and RK-SE methods developed in this paper, as well as the variance dissipation method used to assess numerical mixing and close tracer budgets, are available in Oceananigans, starting from version 0.104. The simulations shown in this paper, which also showcase how to use the above methods, are publicly available (Silvestri, 2025). The above-mentioned repository also includes the implementation of the linear stability analysis for both RK-IM and RK-SE and the notebooks that generate the figures. AI tools were used to assist with proofreading and correcting the manuscript. All AI-generated suggestions were thoroughly reviewed and verified by the authors.

A. Primitive equations in generalized vertical coordinates

Often the numerics benefit from using a vertical coordinate that is different than z and which, in general, can even be moving (vary in time). We denote any such generalized vertical coordinate that evolves with space and time as $r(x, y, z, t)$. The generalized vertical coordinate must vary monotonically with z.

Transforming the equations of motion (including their discrete analogues) in this new generalized coordinate involves a series of chain rule partial derivatives. We have that for any field ϕ:

$$\frac{\partial \phi}{\partial z} = \frac{\partial r}{\partial z}\frac{\partial \phi}{\partial r} \, , \tag{75}$$

$$\frac{\partial \phi}{\partial r} = \frac{\partial z}{\partial r}\frac{\partial \phi}{\partial z} \, . \tag{76}$$

The quantity $\partial z/\partial r$ plays a central role in what follows. We refer to it as "specific thickness" and denote it as σ:

$$\sigma \overset{\text{def}}{=} \left.\frac{\partial z}{\partial r}\right|_{x,y,t} = \left(\left.\frac{\partial r}{\partial z}\right|_{x,y,t}\right)^{-1} \, , \tag{77}$$

where the subscripts next to | denote the quantities that remain constant in the differentiation. The chain rules for differentiation with respect of x, y, or t become:

$$\left.\frac{\partial \phi}{\partial s}\right|_{z} = \left.\frac{\partial \phi}{\partial s}\right|_{r} + \frac{\partial \phi}{\partial r}\frac{\partial r}{\partial s} \, . \tag{78}$$

The horizontal spatial derivatives of the r-coordinate are then rewritten, e.g.,

$$\left.\frac{\partial r}{\partial x}\right|_{y,z,t} = -\left.\frac{\partial z}{\partial x}\right|_{y,r,t}\frac{1}{\sigma} \, . \tag{79}$$

so that the chain rule (78) above becomes

$$\left.\frac{\partial \phi}{\partial x}\right|_{z} = \left.\frac{\partial \phi}{\partial x}\right|_{r} - \frac{1}{\sigma}\frac{\partial \phi}{\partial r}\frac{\partial z}{\partial x} \, . \tag{80}$$

Similarly, we get equivalent expressions for y derivatives as in (79) and (80).

A.1. Continuity equation

With the Boussinesq approximation, the mass conservation reduces to the flow being divergence-less, i.e., $\boldsymbol{\nabla}\cdot\mathbf{u}+\partial_z w=0$. Using the chain rules above, the divergence of the flow in r-coordinates becomes:

$$\begin{aligned}\boldsymbol{\nabla}\cdot\mathbf{u}+\frac{\partial w}{\partial z} &= \left.\frac{\partial u}{\partial x}\right|_z+\left.\frac{\partial v}{\partial y}\right|_z+\frac{\partial w}{\partial z}\\ &= \left.\frac{\partial u}{\partial x}\right|_r+\left.\frac{\partial v}{\partial y}\right|_r-\frac{1}{\sigma}\left(\frac{\partial u}{\partial r}\frac{\partial z}{\partial x}+\frac{\partial v}{\partial r}\frac{\partial z}{\partial y}-\frac{\partial w}{\partial r}\right)\\ &= \frac{1}{\sigma}\left(\left.\frac{\partial\sigma u}{\partial x}\right|_r+\left.\frac{\partial\sigma v}{\partial y}\right|_r-u\left.\frac{\partial\sigma}{\partial x}\right|_r-v\left.\frac{\partial\sigma}{\partial y}\right|_r\right)-\frac{1}{\sigma}\left(\frac{\partial u}{\partial r}\frac{\partial z}{\partial x}+\frac{\partial v}{\partial y}\frac{\partial z}{\partial y}-\frac{\partial w}{\partial r}\right).\end{aligned}\tag{81}$$

We can rewrite, e.g., $\partial_x\sigma|_r=\partial_r(\partial_x z)$ and similarly for the y direction. After a bit of reordering, we get:

$$\boldsymbol{\nabla}\cdot\mathbf{u}+\frac{\partial w}{\partial z}=\frac{1}{\sigma}\left(\left.\frac{\partial\sigma u}{\partial x}\right|_r+\left.\frac{\partial\sigma v}{\partial y}\right|_r\right)-\frac{1}{\sigma}\frac{\partial}{\partial r}\left(u\frac{\partial z}{\partial x}+v\frac{\partial z}{\partial y}-w\right).\tag{82}$$

Note that above, w is the vertical velocity referenced to the z coordinate. The vertical velocity w_r of the r surface referenced to the z coordinate is

$$w_r\overset{\text{def}}{=}\left.\frac{\partial z}{\partial t}\right|_r+u\frac{\partial z}{\partial x}+v\frac{\partial z}{\partial y}.\tag{83}$$

Thus, the vertical velocity across the r surfaces is the difference between w and w_r

$$\begin{aligned}\omega &\overset{\text{def}}{=} w-w_r\\ &= w-\left.\frac{\partial z}{\partial t}\right|_r-u\frac{\partial z}{\partial x}-v\frac{\partial z}{\partial y}.\end{aligned}\tag{84}$$

With the definition of ω in (82) we get

$$\begin{aligned}\boldsymbol{\nabla}\cdot\mathbf{u}+\frac{\partial w}{\partial z} &= \frac{1}{\sigma}\left(\left.\frac{\partial\sigma u}{\partial x}\right|_r+\left.\frac{\partial\sigma v}{\partial y}\right|_r\right)+\frac{1}{\sigma}\frac{\partial}{\partial r}\left(\omega+\left.\frac{\partial z}{\partial t}\right|_r\right)\\ &= \frac{1}{\sigma}\left(\left.\frac{\partial\sigma u}{\partial x}\right|_r+\left.\frac{\partial\sigma v}{\partial y}\right|_r\right)+\frac{1}{\sigma}\frac{\partial\omega}{\partial r}+\frac{1}{\sigma}\frac{\partial\sigma}{\partial t},\end{aligned}\tag{85}$$

which implies that the mass conservation takes the form of (12), i.e.,

$$\frac{\partial\sigma}{\partial t}+\boldsymbol{\nabla}\cdot(\sigma\mathbf{u})+\frac{\partial\omega}{\partial r}=0.\tag{86}$$

A.2. Tracer equation

The evolution equation for a tracer C, which also includes vertical diffusion, reads

$$\left.\frac{\partial C}{\partial t}\right|_z+\boldsymbol{\nabla}\cdot(\mathbf{u}C)+\frac{\partial wC}{\partial z}=\frac{\partial}{\partial z}\left(\kappa\frac{\partial C}{\partial z}\right).\tag{87}$$

Using the same procedure as in A.1 for the continuity equation, the left-hand side of (87) yields:

$$\begin{aligned}&\left.\frac{\partial C}{\partial t}\right|_z+\boldsymbol{\nabla}\cdot(\mathbf{u}C)+\frac{\partial wC}{\partial z}=\\ &\quad=\left.\frac{\partial C}{\partial t}\right|_z+\frac{1}{\sigma}\left(\left.\frac{\partial\sigma uC}{\partial x}\right|_r+\left.\frac{\partial\sigma vC}{\partial y}\right|_r\right)+\frac{1}{\sigma}\frac{\partial}{\partial r}\left(C\omega+C\left.\frac{\partial z}{\partial t}\right|_r\right)\\ &\quad=\left.\frac{\partial C}{\partial t}\right|_z+\frac{1}{\sigma}\left(\left.\frac{\partial\sigma uC}{\partial x}\right|_r+\left.\frac{\partial\sigma vC}{\partial y}\right|_r\right)+\frac{C}{\sigma}\left(\frac{\partial\omega}{\partial r}+\left.\frac{\partial\sigma}{\partial t}\right|_r\right)+\frac{1}{\sigma}\left(\omega+\left.\frac{\partial z}{\partial t}\right|_r\right)\frac{\partial C}{\partial r}\\ &\quad=\left.\frac{\partial C}{\partial t}\right|_z+\frac{1}{\sigma}\left(\left.\frac{\partial\sigma uC}{\partial x}\right|_r+\left.\frac{\partial\sigma vC}{\partial y}\right|_r\right)+\frac{1}{\sigma}\frac{\partial\omega C}{\partial r}+\frac{C}{\sigma}\left.\frac{\partial\sigma}{\partial t}\right|_r+\frac{1}{\sigma}\left.\frac{\partial z}{\partial t}\right|_r\frac{\partial C}{\partial r}.\end{aligned}\tag{88}$$

We recover the time derivative of the tracer at constant r by rewriting the last term using the chain rule for time derivatives:

$$\frac{1}{\sigma}\frac{\partial z}{\partial t}\bigg|_r \frac{\partial C}{\partial r} = \frac{\partial r}{\partial t}\frac{\partial C}{\partial r} = \frac{\partial C}{\partial t}\bigg|_r - \frac{\partial C}{\partial t}\bigg|_z . \tag{89}$$

As such, the left-hand-side of (87) can be rewritten in r-coordinates as

$$\frac{1}{\sigma}\frac{\partial \sigma C}{\partial t}\bigg|_r + \frac{1}{\sigma}\left(\frac{\partial \sigma u C}{\partial x}\bigg|_r + \frac{\partial \sigma v C}{\partial y}\bigg|_r\right) + \frac{1}{\sigma}\frac{\partial \omega C}{\partial r} . \tag{90}$$

Including the vertical diffusion to the right-hand side, and multiplying by σ, we recover the tracer equation (13), i.e.,

$$\frac{\partial \sigma C}{\partial t} + \boldsymbol{\nabla}\cdot(\sigma \mathbf{u} C) + \frac{\partial \omega C}{\partial r} = \frac{\partial}{\partial r}\left(\kappa \frac{\partial C}{\partial z}\right) . \tag{91}$$

A.3. Momentum equations in vector invariant form

The horizontal momentum equation under the hydrostatic approximation reads

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}\bigg|_z + f\hat{\mathbf{z}}\times\mathbf{u} = -\boldsymbol{\nabla} p|_z - g\,\boldsymbol{\nabla}\eta|_z + \frac{\partial}{\partial z}\left(\nu\frac{\partial \mathbf{u}}{\partial z}\right) , \tag{92}$$

where $\mathrm{D}/\mathrm{D}t|_z \overset{\text{def}}{=} \partial_t + \mathbf{u}\cdot\boldsymbol{\nabla} + w\partial_z$ is the material derivative. The above is complemented by the hydrostatic relation

$$\frac{\partial p}{\partial z} = b . \tag{93}$$

Of the above, the Coriolis term is independent of the vertical frame of reference, and the viscous stress is treated similarly to the diffusion of a tracer. In this derivation, we focus on the hydrostatic relation (93) and the material derivative in the momentum equation (92). All the other terms are trivial to derive given the chain rules.

Using the definition (75) of the z-derivative in r-coordinates

$$\frac{\partial p}{\partial r} = \sigma b . \tag{94}$$

The vector-invariant form of the horizontal advection operator is:

$$(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} = \zeta\hat{\mathbf{z}}\times\mathbf{u} + \boldsymbol{\nabla}K , \tag{95}$$

where $\zeta \overset{\text{def}}{=} \partial_x v - \partial_y u$ is the vertical vorticity and $K \overset{\text{def}}{=} |\mathbf{u}|^2/2 = (u^2+v^2)/2$ is the horizontal kinetic energy, such that the material derivative of the horizontal velocity in vector-invariant form is

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}\bigg|_z = \frac{\partial \mathbf{u}}{\partial t}\bigg|_z + \zeta|_z\hat{\mathbf{z}}\times\mathbf{u} + \boldsymbol{\nabla}K|_z + w\frac{\partial \mathbf{u}}{\partial z} . \tag{96}$$

Here, we focus on the u-component of the velocity; the derivation of the v-component follows the same steps. Using the same chain rules as for the tracer equation

$$\frac{\mathrm{D}u}{\mathrm{D}t}\bigg|_z = \frac{\partial u}{\partial t}\bigg|_z - \left(\frac{\partial v}{\partial x}\bigg|_z - \frac{\partial u}{\partial y}\bigg|_z\right)v + \frac{\partial K}{\partial x}\bigg|_z + w\frac{\partial u}{\partial z} \tag{97}$$

$$= \frac{\partial u}{\partial t}\bigg|_z - \Bigg(\underbrace{\frac{\partial v}{\partial x}\bigg|_r - \frac{\partial u}{\partial y}\bigg|_r}_{=\zeta|r} - \frac{1}{\sigma}\frac{\partial v}{\partial r}\frac{\partial z}{\partial x} + \frac{1}{\sigma}\frac{\partial u}{\partial r}\frac{\partial z}{\partial y}\Bigg)v + \frac{\partial K}{\partial x}\bigg|_r - \frac{1}{\sigma}\frac{\partial K}{\partial r}\frac{\partial z}{\partial x} + w\frac{\partial u}{\partial z} . \tag{98}$$

Above, we utilized (78) and (79) repeatedly, e.g., for $\partial_y u|_z$, $\partial_x v|_z$, and $\partial_x K|_z$. Further expanding $\partial_r K = u\partial_r u + v\partial_r v$, a few terms cancel out, and we end up with:

$$\frac{\mathrm{D}u}{\mathrm{D}t}\bigg|_z = \frac{\partial u}{\partial t}\bigg|_z - \zeta|_r v + \frac{\partial K}{\partial x}\bigg|_r + \frac{1}{\sigma}\left(w - v\frac{\partial z}{\partial y} - u\frac{\partial z}{\partial x}\right)\frac{\partial u}{\partial r} . \tag{99}$$

Using the definition of ω in (84), we can rewrite (99) as

$$\left.\frac{\mathrm{D}u}{\mathrm{D}t}\right|_z = \left.\frac{\partial u}{\partial t}\right|_z - \zeta|_r v + \left.\frac{\partial K}{\partial x}\right|_r + \frac{1}{\sigma}\left(\omega + \left.\frac{\partial z}{\partial t}\right|_r\right)\frac{\partial u}{\partial r}\,. \tag{100}$$

Similarly to the tracer derivation in subsection A.2, the last term on the right-hand side, using the chain rule for the time derivative, yields

$$\frac{1}{\sigma}\left(\omega + \left.\frac{\partial z}{\partial t}\right|_r\right)\frac{\partial u}{\partial r} = \frac{\omega}{\sigma}\frac{\partial u}{\partial r} + \left.\frac{\partial u}{\partial t}\right|_r - \left.\frac{\partial u}{\partial t}\right|_z\,, \tag{101}$$

which implies:

$$\left.\frac{\mathrm{D}u}{\mathrm{D}t}\right|_z = \left.\frac{\partial u}{\partial t}\right|_r - \zeta|_r v + \left.\frac{\partial K}{\partial x}\right|_r + \frac{\omega}{\sigma}\frac{\partial u}{\partial r}\,. \tag{102}$$

B. Linear stability analysis: evolution matrices

This appendix provides the detailed derivation of the evolution matrices for both RK-IM and RK-SE used in the linear stability analysis of section 4. The derivation follows loosely the analysis of Demange et al. (2019), where we decompose the variables into vertical modes. To simplify the system, we assume a two-dimensional domain (x, z) in hydrostatic balance, neglect self-advection of momentum, and reduce the advection of buoyancy to wN^2, where $N^2(z)$ is a background stratification. We also linearize the free-surface equations around the rest state and drop the moving vertical coordinate in favor of a fixed coordinate ($\sigma = 1$). Note that, because of the stretching in a moving coordinate, the stability should be easier to achieve than for a fixed coordinate. To derive the modal decomposition, consider the linearized set of coupled equations:

$$\frac{\partial u}{\partial t} = -\frac{\partial p}{\partial x}\,, \tag{103}$$

$$\frac{\partial p}{\partial z} = b\,, \tag{104}$$

$$\frac{\partial b}{\partial t} + N^2 w = 0\,, \tag{105}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0\,, \tag{106}$$

and with boundary conditions including a linearized free surface

$$p = g\eta \ \text{ at } \ z = 0\,, \tag{107}$$

$$\frac{\partial \eta}{\partial t} = w \ \text{ at } \ z = 0\,, \tag{108}$$

$$w = 0 \ \text{ at } \ z = -H\,. \tag{109}$$

By rearranging (103)–(106) we obtain a single equation for pressure:

$$-\frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial}{\partial z}\frac{\partial^2 p}{\partial t^2}\right) = \frac{\partial^2 p}{\partial x^2}\,. \tag{110}$$

By assuming that the solutions are separable between the horizontal and the vertical directions, e.g.,

$$p(x,z,t) = \sum_{q=0}^{\infty} p_q(x,t)\,\phi_q(z)\,, \tag{111}$$

and using (107)–(109) we obtain the following Sturm–Liouville eigenvalue problem

$$-\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{N^2}\frac{\mathrm{d}\phi_q}{\mathrm{d}z}\right) = \frac{1}{c_q^2}\phi_q\,, \tag{112}$$

$$\frac{\mathrm{d}\phi_q}{\mathrm{d}z} = -\frac{N^2}{g}\phi_q \quad \text{at } z = 0\,, \tag{113}$$

$$\frac{\mathrm{d}\phi_q}{\mathrm{d}z} = 0 \quad \text{at } z = -H\,, \tag{114}$$

where $1/c_q^2$ are the eigenvalues and ϕ_q the eigenmodes. To reconstruct pressure p (or any other field), we note that p_q must additionally satisfy the wave equation:

$$\frac{\partial^2 p_q}{\partial x^2} = \frac{1}{c_q^2}\frac{\partial^2 p_q}{\partial t^2} . \tag{115}$$

Assuming a constant background stratification, from (112) and (114), the eigenmodes are

$$\phi_q(z) \sim \cos\left[N(z+H)/c_q\right] , \tag{116}$$

and further enforcing the boundary condition (113) yields a transcendental equation for c_q:

$$\tan\left(\frac{NH}{c_q}\right) = \frac{N}{g} c_q . \tag{117}$$

The modes ϕ_q are orthonormal, i.e., $\int_{-H}^{0} \phi_q \phi_r \,\mathrm{d}z = \delta_{qr}$. Under the depth-independent barotropic mode assumption, the barotropic mode, ϕ_0, can be normalized so that $\int_{-H}^{0} \phi_0 \,\mathrm{d}z \approx H$, while all other modes integrate to 0, i.e., $\int_{-H}^{0} \phi_q \,\mathrm{d}z \approx 0$ for $q \geq 1$.

We want to derive the evolution matrices of the coupled barotropic–baroclinic system for both the implicit and split–explicit free surface discretizations. For this reason, we apply a common set of simplifying assumptions to proceed in the analysis. First of all, as done by Demange et al. (2019), we truncate to the first two modes, representing barotropic evolution (mode 0) and the fastest baroclinic mode (mode 1). Furthermore, we simplify the eigenvalues, where, to first order, we can assume that

$$c_0 \approx \sqrt{gH} , \quad c_1 \approx NH/\pi . \tag{118}$$

For notational convenience, we denote $\phi_1(0) = \psi$ and let $\phi_0(0) = 1$. In principle, the depth-independent assumption and the eigenvalue simplification are only valid in the limit $NH \ll \sqrt{gH}$, which indeed holds for the ocean. Using these simplifying assumptions, we rewrite velocity, pressure, and free surface

$$u(x,z,t) = u_0(x,t) + u_1(x,t)\phi_1(z) , \tag{119}$$

$$\eta(x,t) = h_0(x,t) + h_1(x,t)\psi_1 , \tag{120}$$

$$p(x,z,t) = gh_1(x,t)\left[\phi_1(z) - \psi_1\right] , \tag{121}$$

such that

$$p + g\eta = gh_0(x,t) + gh_1(x,t)\phi_1(z) . \tag{122}$$

The vertical velocity w is obtained by the continuity equation and buoyancy from the hydrostatic relation

$$w(x,z,t) \approx -(z+H)\frac{\partial u_0}{\partial x} - \Phi_1(z)\frac{\partial u_1}{\partial x} , \tag{123}$$

$$b(x,z,t) \approx gh_1 \frac{\mathrm{d}\phi_1}{\mathrm{d}z} , \tag{124}$$

where $\Phi(z) \overset{\text{def}}{=} \int_{-H}^{z} \phi(z')\,\mathrm{d}z'$. Note that an individual RK substep evolves as

$$\hat{y}^{m+1} = \mathbf{A}_m \hat{y}^n + \mathbf{B}_m \hat{y}^m , \tag{125}$$

where $\hat{y}$ are the Fourier components of the decomposed state, and the matrices $\mathbf{A}$ and $\mathbf{B}$ depend on the timestep of the RK substep (γ). The substeps need to be combined to obtain the evolution matrix of the complete RK scheme, which is used to compute the amplification factor. The complete evolution matrix can be built recursively from the substep matrices as

$$\mathbf{M}_M = \mathbf{A}_{M-1} + \mathbf{B}_{M-1}\,\mathbf{M}_{M-1} , \qquad \mathbf{M}_0 = \mathbf{I} . \tag{126}$$

B.1. Implicit free surface

We start from the already discretized primitive equations, by decomposing into barotropic and baroclinic mode equations (53)–(56). We first derive the evolution matrices for $w^p = w^m$. For the predictor velocity equation (53), substituting the modal decomposition for p^m and u^m, yields:

$$u_0^* + u_1^*\phi_1 = u_0^n + u_1^n\phi_1 - \gamma g \frac{\partial h_1^m}{\partial x}(\phi_1 - \psi)\,, \tag{127}$$

Projecting onto modes ϕ_q and integrating in depth

$$u_0^* = u_0^n + \gamma g \frac{\partial h_1^m}{\partial x}\psi\,, \tag{128}$$

$$u_1^* = u_1^n - \gamma g \frac{\partial h_1^m}{\partial x}\,. \tag{129}$$

The extra coupling term on the right-hand side of u_0^* represents the “predictor-corrector” algorithm of the implicit free surface, where the predictor velocity u^* is advanced using only the internal pressure absent the free surface. The modal decomposition of the buoyancy equation (54), substituting b and w modes

$$h_1^{m+1}\frac{\mathrm{d}\phi_1}{\mathrm{dz}} = h_1^n\frac{\mathrm{d}\phi_1}{\mathrm{dz}} + \frac{\gamma}{g}N^2\left[(z+H)\frac{\partial u_0^m}{\partial x} + \Phi_1(z)\frac{\partial u_1^m}{\partial x}\right]\,, \tag{130}$$

which, after a vertical integration, returns an equation for h_1^{m+1}

$$h_1^{m+1} = h_1^n - \frac{\gamma c_1^2}{g}\left(\frac{\partial u_1^m}{\partial x} - \psi_2\frac{\partial u_0^m}{\partial x}\right)\,, \tag{131}$$

where $\psi_2 = \pi^2(4\psi)^{-1}$. Moving on to the free surface evolution (55),

$$h_0^{m+1} + h_1^{m+1}\psi - \gamma^2 gH\left(\frac{\partial^2 h_0^{m+1}}{\partial x^2} + \frac{\partial^2 h_1^{m+1}}{\partial x^2}\psi\right) = h_0^n + h_1^n\psi - \gamma H\frac{\partial u_0^*}{\partial x}\,, \tag{132}$$

while the final velocity update is derived as

$$u_0^{m+1} = u_0^* - \gamma g\left(\frac{\partial h_0^{m+1}}{\partial x} + \frac{\partial h_1^{m+1}}{\partial x}\psi\right)\,, \tag{133}$$

$$u_1^{m+1} = u_1^*\,. \tag{134}$$

Note that, in the above equations, the coupling in the velocity evolution equations arises from decomposing the pressure into an internal hydrostatic component, p, that vanishes at $z = 0$ and a surface component, $g\eta$, a typical approximation used in hydrostatic ocean models. The final decomposed equations read:

$$u_0^* = u_0^n + \gamma g\psi\frac{\partial h_1^m}{\partial x}\,, \tag{135}$$

$$u_1^{m+1} = u_1^n - \gamma g\frac{\partial h_1^m}{\partial x}\,, \tag{136}$$

$$h_1^{m+1} = h_1^n - \gamma\frac{c_1^2}{g}\left(\frac{\partial u_1^m}{\partial x} - \psi_2\frac{\partial u_0^m}{\partial x}\right)\,, \tag{137}$$

$$h_0^{m+1} - \gamma^2 gH\frac{\partial^2 h_0^{m+1}}{\partial x^2} = -h_1^{m+1}\psi + \gamma^2 gH\psi\frac{\partial^2 h_1^{m+1}}{\partial x^2} + h_0^n + h_1^n\psi - \gamma H\frac{\partial u_0^*}{\partial x}\,, \tag{138}$$

$$u_0^{m+1} = u_0^* - \gamma g\frac{\partial h_0^{m+1}}{\partial x} - \gamma g\psi\frac{\partial h_1^{m+1}}{\partial x}\,. \tag{139}$$

To lighten the notation, it is convenient to define

$$D \overset{\text{def}}{=} 1 + \gamma^2 gHk^2 \quad \text{and} \quad C \overset{\text{def}}{=} \gamma\frac{c_1^2}{g}\,, \tag{140}$$

where D is the barotropic damping coefficient that introduces numerical dissipation in the barotropic (and via the coupling also the baroclinic) system. We complete the derivation with a Fourier substitution for the time-varying horizontal modes, e.g.,

$$u_q(x,t) = \hat{u}_q(t)\,e^{ikx}\,, \tag{141}$$

and similarly for other fields, e.g., h, p, etc. When we apply (141) to the baroclinic system of equations (136)–(137) we get:

$$\hat{u}_1^{m+1} = \hat{u}_1^n - \gamma g i k \hat{h}_1^m \,, \tag{142}$$

$$\hat{h}_1^{m+1} = \hat{h}_1^n - ikC(\hat{u}_1^m - \psi_2 \hat{u}_0^m) \,. \tag{143}$$

For the barotropic mode, we solve the free surface equation (138) using $\hat{h}_1^{m+1}$ as computed from the buoyancy equation.

$$\hat{u}_0^* = \hat{u}_0^n + \gamma g \psi i k \hat{h}_1^m \,, \tag{144}$$

$$\hat{h}_0^{m+1} D = \hat{h}_0^n + \hat{h}_1^n \psi - \gamma H i k \hat{u}_0^* - \hat{h}_1^{m+1} \psi D \,, \tag{145}$$

$$\hat{u}_0^{m+1} = \hat{u}_0^* - \gamma g i k \left(\hat{h}_0^{m+1} + \psi \hat{h}_1^{m+1} \right) \,. \tag{146}$$

The above equations illustrate the predictor-corrector algorithm acting on the barotropic mode. The explicit evolution of $\hat{h}_0^{m+1}$ is then given by

$$\hat{h}_0^{m+1} = \frac{\hat{h}_0^n + \psi \hat{h}_1^n - \gamma H i k \hat{u}_0^*}{D} - \psi \hat{h}_1^{m+1} \,. \tag{147}$$

Substituting for $\hat{u}_0^*$ and $\hat{h}_1^{m+1}$ yields

$$\hat{h}_0^{m+1} = \frac{\hat{h}_0^n + \psi \hat{h}_1^n}{D} - \frac{\gamma H i k}{D} \left(\hat{u}_0^n + \gamma g \psi i k \hat{h}_1^m \right) - \psi \left(\hat{h}_1^n - ikC \hat{u}_1^m + ikC \psi_2 \hat{u}_0^m \right) \,, \tag{148}$$

which, after rearranging the terms, becomes

$$\hat{h}_0^{m+1} = -\frac{\gamma H i k}{D} \hat{u}_0^n + \frac{\hat{h}_0^n}{D} - \hat{h}_1^n \frac{\psi(D-1)}{D} + \hat{h}_1^m \frac{\psi(D-1)}{D} + \psi i k C \left(\hat{u}_1^m - \psi_2 \hat{u}_0^m \right) \,. \tag{149}$$

Finally, the evolution of the barotropic velocity component $\hat{u}_0^{m+1}$, after substitution with $\hat{u}_0^*$, $\hat{h}_0^{m+1}$, and $\hat{h}_1^{m+1}$ reads

$$\hat{u}_0^{m+1} = \frac{\hat{u}_0^n}{D} - \frac{\gamma g i k}{D} \hat{h}_0^n - \frac{\gamma g i k \psi}{D} \hat{h}_1^n + \frac{\gamma g i k \psi}{D} \hat{h}_1^m \,, \tag{150}$$

where the dependencies of $\hat{u}_0^{m+1}$ on the substepped $\hat{u}_0^m$ and $\hat{u}_1^m$ disappear. Therefore, the evolution of the barotropic component of the velocity depends only on the previous state (n) and the evolved state of the baroclinic height $\hat{h}_1^m$ in accordance with the implicit free surface discretization. With these equations in place, we can define the evolution matrices for one RK-IM substep where the solution vector is $\hat{y} = [u_0, h_0, u_1, h_1]^T$, as

$$\mathbf{A}_{\text{RK-IM}} = \begin{pmatrix} \frac{1}{D} & -\frac{\gamma g i k}{D} & 0 & -\frac{\gamma g i k \psi}{D} \\ -\frac{\gamma H i k}{D} & \frac{1}{D} & 0 & -\frac{\psi(D-1)}{D} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{151}$$

$$\mathbf{B}_{\text{RK-IM}} = \begin{pmatrix} 0 & 0 & 0 & \frac{\gamma g i k \psi}{D} \\ -ikC\psi\psi_2 & 0 & ikC\psi & \frac{\psi(D-1)}{D} \\ 0 & 0 & 0 & -\gamma g i k \\ ikC\psi_2 & 0 & -ikC & 0 \end{pmatrix} . \tag{152}$$

We now compute the evolution matrices for $w^p = w^\dagger$ where

$$w^\dagger = -(z+H) \frac{\partial u_0^{m+1}}{\partial x} - \Phi_1(z) \frac{\partial u_1^m}{\partial x} \,. \tag{153}$$

The matrix rows corresponding to the barotropic and baroclinic velocity evolution equations do not change. With respect to the evolution of h_1^{m+1},

$$\hat{h}_1^{m+1} = \hat{h}_1^n - ikC(\hat{u}_1^m - \psi_2 \hat{u}_0^{m+1}) \,, \tag{154}$$

which, after some algebra, simplifies to

$$\hat{h}_1^{m+1} = \left(1 + \frac{D_2 - 1}{D}\right)\hat{h}_1^n - ikC\hat{u}_1^m + \frac{ikC\psi_2}{D}\hat{u}_0^n + \frac{\gamma g k^2 C\psi_2}{D}\hat{h}_0^n - \frac{D_2 - 1}{D}\hat{h}_1^m \,. \tag{155}$$

For notational convenience, we have denoted

$$D_2 \overset{\text{def}}{=} 1 + \gamma g k^2 C \psi_2 \psi \,. \tag{156}$$

Finally, the equation for $\hat{h}_0^{m+1}$ is obtained by (147) with substitution of the above equation for h_1^{m+1},

$$\hat{h}_0^{m+1} = \frac{2 - D_2}{D}\hat{h}_0^n - ik\frac{\gamma H + C\psi_2\psi}{D}\hat{u}_0^n + \psi\frac{2 - D - D_2}{D}\hat{h}_1^n + ikC\psi\hat{u}_1^m + \psi\frac{D + D_2 - 2}{D}\hat{h}_1^m \tag{157}$$

The new evolution matrices with $w^p = w^\dagger$ read

$$\mathbf{A}_{\text{RK-IM}} = \begin{pmatrix} \frac{1}{D} & -\frac{\gamma g i k}{D} & 0 & -\frac{\gamma g i k \psi}{D} \\ -ik\frac{\gamma H + C\psi_2\psi}{D} & \frac{2 - D_2}{D} & 0 & \frac{\psi(2 - D_2 - D)}{D} \\ 0 & 0 & 1 & 0 \\ \frac{ikC\psi_2}{D} & \frac{\gamma g k^2 C\psi_2}{D} & 0 & 1 + \frac{D_2 - 1}{D} \end{pmatrix}, \tag{158}$$

$$\mathbf{B}_{\text{RK-IM}} = \begin{pmatrix} 0 & 0 & 0 & \frac{\gamma g i k \psi}{D} \\ 0 & 0 & ikC\psi & \frac{\psi(D + D_2 - 2)}{D} \\ 0 & 0 & 0 & -\gamma g i k \\ 0 & 0 & -ikC & \frac{1 - D_2}{D} \end{pmatrix}. \tag{159}$$

The one noticeable difference is the absence of the stage barotropic components u_0^m and h_0^m in the **B** evolution matrix, dropped in favor of new coupling terms in the **A** matrix which includes contributions from the previous time step (u_0^n and h_0^n).

B.2. Split–explicit free surface

The modal decomposition of the baroclinic system, with $w^p = w^m$, has the familiar form given by (136)–(137). For the barotropic system, we note that $U^n = Hu_0^n$, so we introduce a fast barotropic velocity $u_0^s \overset{\text{def}}{=} U^s/H$. η^s is an auxiliary barotropic variables to be averaged, so we avoid substituting it with h_q. The modal decomposition of the barotropic step, with the same assumptions used in the previous subsection, leads to

$$u_0^{s+1} = u_0^s - \gamma_B g\left(\frac{\partial\eta^s}{\partial x} - \psi\frac{\partial h_1^m}{\partial x}\right), \tag{160}$$

$$\eta^{s+1} = \eta^s - \gamma_B H\frac{\partial u_0^{s+1}}{\partial x}, \tag{161}$$

$$u_0^{m+1} = \sum_{s=1}^{N_S} \alpha^s u_0^s \,, \tag{162}$$

$$h_0^{m+1} = \sum_{s=1}^{N_S} \alpha^s \eta^s - \psi h_1^{m+1} \,. \tag{163}$$

Focusing on the barotropic sub-cycling procedure and transforming the barotropic variables into their respective coefficients, we get

$$\hat{u}_0^{s+1} = \hat{u}_0^s - \gamma_B i k g\left(\hat{\eta}^s - \psi\hat{h}_1^m\right), \tag{164}$$

$$\hat{\eta}^{s+1} = \hat{\eta}^s - \gamma_B i k H\hat{u}_0^{s+1} \,. \tag{165}$$

Therefore, the barotropic system evolves according to

$$\begin{pmatrix} \hat{u}_0^{s+1} \\ \hat{\eta}^{s+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -\gamma_B i k g \\ -\gamma_B i k H & 1 - \gamma_B^2 k^2 g H \end{pmatrix}}_{\mathbf{C}} \begin{pmatrix} \hat{u}_0^s \\ \hat{\eta}^s \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & \gamma_B i k g \psi \\ 0 & \gamma_B^2 k^2 g H \psi \end{pmatrix}}_{\mathbf{D}} \begin{pmatrix} 0 \\ \hat{h}_1^m \end{pmatrix}. \tag{166}$$

Where **C** expresses the forward-backward sub-cycling while **D** accounts for the “frozen” hydrostatic pressure gradient. Since $\hat{\eta}^{s=0} = \hat{\eta}^n = \hat{h}_0^n + \psi \hat{h}_1^n$ and $\hat{u}^{s=0} = \hat{u}_0^n$, after $s+1$ substeps, we get that

$$\begin{pmatrix} \hat{u}_0^{s+1} \\ \hat{\eta}^{s+1} \end{pmatrix} = [\mathbf{C}]^{s+1} \begin{pmatrix} \hat{u}_0^n \\ \hat{h}_0^n + \psi \hat{h}_1^n \end{pmatrix} + \left(\sum_{s'=0}^{s} [\mathbf{C}]^{s'} \, \mathbf{D} \right) \begin{pmatrix} 0 \\ \hat{h}_1^m \end{pmatrix} . \tag{167}$$

To compute values at $m+1$, we apply the filter weights and subtract the updated baroclinic height

$$\begin{pmatrix} \hat{u}_0^{m+1} \\ \hat{h}_0^{m+1} \end{pmatrix} = \mathbf{A}_{BT} \begin{pmatrix} \hat{u}_0^n \\ \hat{h}_0^n \end{pmatrix} + \psi \mathbf{A}_{BT} \begin{pmatrix} 0 \\ \hat{h}_1^n \end{pmatrix} + \mathbf{B}_{BT} \begin{pmatrix} 0 \\ \hat{h}_1^m \end{pmatrix} + \mathbf{P} \begin{pmatrix} 0 \\ \hat{h}_1^{m+1} \end{pmatrix} . \tag{168}$$

where the barotropic matrices read

$$\mathbf{A}_{BT} = \sum_{s=1}^{N_S} \alpha^s \, [\mathbf{C}]^s \ , \quad \mathbf{B}_{BT} = \sum_{s=1}^{N_S} \alpha^s \left(\sum_{s'=1}^{s} [\mathbf{C}]^{s'-1} \, \mathbf{D} \right) , \quad \mathbf{P} = \begin{pmatrix} 0 & 0 \\ 0 & -\psi \end{pmatrix} . \tag{169}$$

Rewriting $\hat{h}_1^{m+1}$ using (143) gives

$$\begin{pmatrix} \hat{u}_0^{m+1} \\ \hat{h}_0^{m+1} \end{pmatrix} = \mathbf{A}_{BT} \begin{pmatrix} \hat{u}_0^n \\ \hat{h}_0^n \end{pmatrix} + (\psi \mathbf{A}_{BT} + \mathbf{P}) \begin{pmatrix} 0 \\ \hat{h}_1^n \end{pmatrix} + \mathbf{B}_{BT} \begin{pmatrix} 0 \\ \hat{h}_1^m \end{pmatrix} - ikC\mathbf{P} \begin{pmatrix} 0 \\ \hat{u}_1^m - \psi_2 \hat{u}_0^m \end{pmatrix} . \tag{170}$$

We write the two evolution matrices **A** and **B** of an RK-SE substep as functions of the coefficients of the barotropic matrices $\mathbf{A}_{BT}$ and $\mathbf{B}_{BT}$, which we indicate with a and b:

$$\mathbf{A}_{\text{RK-SE}} = \begin{pmatrix} a_{11} & a_{12} & 0 & \psi a_{12} \\ a_{21} & a_{22} & 0 & \psi(a_{22} - 1) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{171}$$

$$\mathbf{B}_{\text{RK-SE}} = \begin{pmatrix} 0 & 0 & 0 & b_{12} \\ -ikC\psi\psi_2 & 0 & ikC\psi & b_{22} \\ 0 & 0 & 0 & -\gamma gik \\ ikC\psi_2 & 0 & -ikC & 0 \end{pmatrix} . \tag{172}$$

The baroclinic part of the matrices derived for the RK-SE discretization remains the same as for RK-IM, since the time discretization of the baroclinic component did not change. The barotropic structure of **B** also remains very similar, where the damping is provided by the filtering weights of the barotropic sub-cycling rather than the implicit factor D.

If using $w^p = w^\dagger$ as in equation (64), the coupling is slightly more complicated as information from the barotropic sub-cycling enters the buoyancy equation directly. The decomposed w^p follows

$$w^p = -(z+H) \frac{\partial u_0^\dagger}{\partial x} - \Phi_1(z) \frac{\partial u_1^\dagger}{\partial x} \ , \quad \text{with} \quad u_1^\dagger = u_1^m \, . \tag{173}$$

The baroclinic height evolution in Fourier space now includes the updated $u_0^\dagger$, where

$$\hat{h}_1^{m+1} = \hat{h}_1^n - ikC \left(\hat{u}_1^m - \psi_2 \hat{u}_0^\dagger \right) . \tag{174}$$

Since $u_0^\dagger = \sum_{s=1}^{N_S} \beta^s U^s$, we can compute its Fourier coefficient, in analogy to $\hat{u}_0^{m+1}$, through equation (168)

$$\hat{u}_0^\dagger = \tilde{a}_{11} \hat{u}_0^n + \tilde{a}_{12} \hat{h}_0^n + \psi \tilde{a}_{12} \hat{h}_1^n + \tilde{b}_{12} \hat{h}_1^m \, , \tag{175}$$

where $\tilde{a}$ and $\tilde{b}$ are coefficients of complementary barotropic matrices computed using the β^s transport weights

$$\tilde{\mathbf{A}}_{BT} = \sum_{s=1}^{N_S} \beta^s \, [\mathbf{C}]^s \ , \quad \tilde{\mathbf{B}}_{BT} = \sum_{s=1}^{N_S} \beta^s \left(\sum_{s'=1}^{s} [\mathbf{C}]^{s'-1} \, \mathbf{D} \right) , \tag{176}$$

The evolution of the baroclinic height then becomes

$$\hat{h}_1^{m+1} = \hat{h}_1^n - ikC \hat{u}_1^m + ikC\psi_2 \left(\tilde{a}_{11} \hat{u}_0^n + \tilde{a}_{12} \hat{h}_0^n + \psi \tilde{a}_{12} \hat{h}_1^n + \tilde{b}_{12} \hat{h}_1^m \right) , \tag{177}$$

and the barotropic height evolution is modified to include this new transport velocity

$$\begin{aligned}\hat{h}_0^{m+1} = a_{21}\hat{u}_0^n + a_{22}\hat{h}_0^n + \psi(a_{22}-1)\hat{h}_1^n \\ - ikC\psi\psi_2\left(\tilde{a}_{11}\hat{u}_0^n + \tilde{a}_{12}\hat{h}_0^n + \psi\tilde{a}_{12}\hat{h}_1^n + \tilde{b}_{12}\hat{h}_1^m\right) \\ + ikC\psi\hat{u}_1^m + b_{22}\hat{h}_1^m \,.\end{aligned} \tag{178}$$

The coupling now becomes much tighter, and the matrices are modified to

$$\mathbf{A}_{\text{RK-SE}} = \begin{pmatrix} a_{11} & a_{12} & 0 & \psi a_{12} \\ a_{21} - ikC\psi\psi_2\tilde{a}_{11} & a_{22} - ikC\psi\psi_2\tilde{a}_{12} & 0 & \psi(a_{22}-1) - ikC\psi^2\psi_2\tilde{a}_{12} \\ 0 & 0 & 1 & 0 \\ ikC\psi_2\tilde{a}_{11} & ikC\psi_2\tilde{a}_{12} & 0 & 1 + ikC\psi\psi_2\tilde{a}_{12} \end{pmatrix}, \tag{179}$$

$$\mathbf{B}_{\text{RK-SE}} = \begin{pmatrix} 0 & 0 & 0 & b_{12} \\ 0 & 0 & ikC\psi & b_{22} - ikC\psi\psi_2\tilde{b}_{12} \\ 0 & 0 & 0 & -\gamma gik \\ 0 & 0 & -ikC & ikC\psi_2\tilde{b}_{12} \end{pmatrix}. \tag{180}$$

The structure of **A** and **B** using this vertical velocity definition differs from the previously derived matrices, as, similarly to the changes in the implicit free surface evolution, the dependencies on the barotropic components at stage *m* are dropped in favor of extra coupling terms depending on the previous state *n*.

C. Assessing numerical mixing

Having a robust method to quantify the tracer variance budget is necessary to judge and compare different numerical discretizations among each other. Various methods have been developed in the ocean modeling literature (Klingbeil et al., 2014; Henell et al., 2023; Banerjee et al., 2024; Schifano et al., 2025). These methods were developed for single-step, single-stage methods and require adaptation to be used with the time discretization approaches we analyze in this paper. In this section, we use the exact methodology described by Banerjee et al. (2024), which is the only methodology that can be adapted to the AB2 scheme currently implemented in Oceananigans, and to the RK scheme developed in this paper. For the latter, the low-storage formulation simplifies the derivation as only the last substep contributes to the buoyancy budget, and it is the only one that needs to be accounted for to quantify the numerical mixing. Therefore, the methodology described by Banerjee et al. (2024) reads

$$P_x = 2F_\star^x \delta_i \left(\frac{T^{n+1} + T^n}{2}\right) - U_\star \delta_i \left(T^{n+1}T^n\right) , \tag{181}$$

$$P_y = 2F_\star^y \delta_j \left(\frac{T^{n+1} + T^n}{2}\right) - V_\star \delta_j \left(T^{n+1}T^n\right) , \tag{182}$$

$$P_z = 2F_\star^z \delta_k \left(\frac{T^{n+1} + T^n}{2}\right) - W_\star \delta_k \left(T^{n+1}T^n\right) , \tag{183}$$

where δ_d denotes a difference in the d direction. For the RK discretization:

$$F_\star^x = \mathcal{A}_{M-1}^x u_{M-1}\widetilde{T}_i^n \quad \text{and} \quad U_\star = \mathcal{A}_{M-1}^x u_{M-1} \,, \tag{184}$$

$$F_\star^y = \mathcal{A}_{M-1}^y v_{M-1}\widetilde{T}_j^n \quad \text{and} \quad V_\star = \mathcal{A}_{M-1}^y v_{M-1} \,, \tag{185}$$

$$F_\star^z = \mathcal{A}_{M-1}^z w_{M-1}\widetilde{T}_k^n \quad \text{and} \quad W_\star = \mathcal{A}_{M-1}^z w_{M-1} \,, \tag{186}$$

where the subscript indicates the substep $(M-1)$, $\mathcal{A}$ are cell-areas, while $\widetilde{T}_i$, $\widetilde{T}_j$, and $\widetilde{T}_k$ indicate the finite volume reconstructions performed on i, j, and k faces, respectively. Whereas for the AB2 time discretization

$$F_\star^x = \mathcal{A}_n^x \left(\frac{3}{2} + \epsilon\right) u^n \widetilde{T}_i^n - \mathcal{A}_{n-1}^x \left(\frac{1}{2} + \epsilon\right) u^{n-1}\widetilde{T}_i^{n-1} \,, \tag{187}$$

$$F_\star^y = \mathcal{A}_n^y \left(\frac{3}{2} + \epsilon\right) v^n \widetilde{T}_j^n - \mathcal{A}_{n-1}^y \left(\frac{1}{2} + \epsilon\right) v^{n-1}\widetilde{T}_j^{n-1} \,, \tag{188}$$

$$F_\star^z = \mathcal{A}_n^z \left(\frac{3}{2} + \epsilon\right) w^n \widetilde{T}_k^n - \mathcal{A}_{n-1}^z \left(\frac{1}{2} + \epsilon\right) w^{n-1}\widetilde{T}_k^{n-1} \,, \tag{189}$$

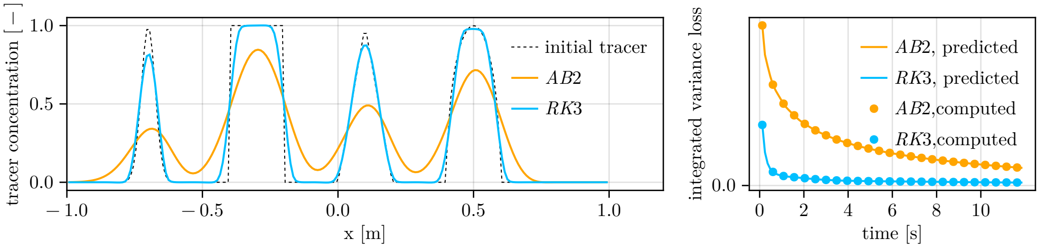

Figure 12: Validation of the tracer numerical mixing approach in a one-dimensional wave-advection test case. The left panel shows the advected tracer for AB2 and RK3 after 6 cycles compared to the initial distribution. The right panel shows the comparison between the integral variance dissipation computed using the dissipation metric and equation (193).

and

$$U_\star = \mathcal{A}^x_n \left(\frac{3}{2} + \epsilon\right) u^n - \mathcal{A}^x_{n-1} \left(\frac{1}{2} + \epsilon\right) u^{n-1} \,, \tag{190}$$

$$V_\star = \mathcal{A}^y_n \left(\frac{3}{2} + \epsilon\right) v^n - \mathcal{A}^y_{n-1} \left(\frac{1}{2} + \epsilon\right) v^{n-1} \,, \tag{191}$$

$$W_\star = \mathcal{A}^z_n \left(\frac{3}{2} + \epsilon\right) w^n - \mathcal{A}^z_{n-1} \left(\frac{1}{2} + \epsilon\right) w^{n-1} \,. \tag{192}$$

Figure 12 shows a validation of the above method applied to the RK3 and AB2 discretization in the simple context of a one-dimensional tracer advection configuration. We advect an initial tracer profile consisting of various shapes by a constant velocity field $u = 1$. We use CFL of 0.2 for AB2 and 0.5 for RK3. The left panel shows the final tracer distribution after 6 cycles compared to the initial tracer distribution, while the right panel shows the comparison of the integrated tracer variance dissipation, computed as

$$\int_{-1}^{1} \frac{\left(c^{n+1}\right)^2 - \left(c^n\right)^2}{\Delta t} \, \mathrm{d}x \,, \tag{193}$$

and the same quantity computed using the technique described in this section. The two different approaches match up to floating-point precision, validating the correct implementation of the method. We use this metric to assess diapycnal mixing. Since we use a linear equation of state that depends both on temperature T and salinity S, we can compute diapycnal mixing, or mixing of buoyancy, by projecting the tracer fluxes in buoyancy space:

$$F_\star(b) = g\alpha F_\star(T) - g\beta F_\star(S) \,, \tag{194}$$

where α and β are the (constant) thermal expansion and haline contraction coefficients.

Acknowledgements:
Your acknowledgements will appear here.

References

Abernathey, R., Marshall, J., & Ferreira, D. (2011). The dependence of Southern Ocean meridional overturning on wind stress. *Journal of Physical Oceanography*, *41*(12), 2261-2278. doi: 10.1175/JPO-D-11-023.1

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., … Zhang, R. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3167-3211. doi: 10.1029/2019MS001726

Adcroft, A., & Campin, J.-M. (2004). Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modelling*, *7*, 269-284. doi: 10.1016/j.ocemod.2003.09.003

Asselin, R. (1972). Frequency filter for time integrations. *Monthly Weather Review*, *100*(6), 487–490. doi: 10.1175/1520-0493(1972)100<0487:FFFTI>2.3.CO;2

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Monthly Weather Review*, *139*(12), 3887–3905. doi: 10.1175/MWR-D-10-05013.1

Banerjee, T., Danilov, S., Klingbeil, K., & Campin, J.-M. (2024). Discrete variance decay analysis of spurious mixing. *Ocean Modelling*, *192*, 102460. doi: 10.1016/j.ocemod.2024.102460

Campin, J.-M., Adcroft, A., Hill, C., & Marshall, J. (2004). Conservation of properties in a free-surface model. *Ocean Modelling*, *6*(3), 221-244. doi: 10.1016/S1463-5003(03)00009-X

Climate Modeling Alliance, & contributors. (2026). *Oceananigans.jl.* Retrieved 2026-04-09, from `https://clima.github.io/OceananigansDocumentation/stable/literated/internal_tide`

Conde, S., Gottlieb, S., Grant, Z., & Shadid, J. (2017). Implicit and implicit–explicit strong stability preserving Runge–Kutta methods with high linear order. *Journal of Scientific Computing*, *73*. doi: 10.1007/s10915-017-0560-2

Danilov, S., Sidorenko, D., Wang, Q., & Jung, T. (2017). The Finite-volumE Sea ice-Ocean Model (FESOM2). Part 1: Overview of the model. *Geoscientific Model Development*, *10*(2), 765-789. doi: 10.5194/gmd-10-765-2017

Demange, J., Debreu, L., Marchesiello, P., Lemarié, F., Blayo, E., & Eldred, C. (2019). Stability analysis of split-explicit free surface ocean models: Implication of the depth-independent barotropic mode approximation. *Journal of Computational Physics*, *398*, 108875. doi: 10.1016/j.jcp.2019.108875

Dukowicz, J. K., & Smith, R. D. (1994). Implicit free-surface method for the Bryan-Cox-Semtner ocean model. *Journal of Geophysical Research: Oceans*, *99*(C4), 7991-8014. doi: 10.1029/93JC03455

Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, *8*(10), 3071–3104. doi: 10.5194/gmd-8-3071-2015

Griffies, S. M., Pacanowski, R. C., Schmidt, M., & Balaji, V. (2001). Tracer conservation with an explicit free surface method for z-coordinate ocean models. *Monthly Weather Review*, *129*(5), 1081-1098. doi: 10.1175/1520-0493(2001)129<1081:TCWAEF>2.0.CO;2

Henell, E., Burchard, H., Gräwe, U., & Klingbeil, K. (2023). Spatial composition of the diahaline overturning circulation in a fjord–type, non–tidal estuarine system. *Journal of Geophysical Research: Oceans*, *128*(12), e2023JC019862. (e2023JC019862 2023JC019862) doi: 10.1029/2023JC019862

Hetland, R. D. (2017). Suppression of baroclinic instabilities in buoyancy-driven flow over sloping bathymetry. *Journal of Physical Oceanography*, *47*(1), 49-68. doi: 10.1175/JPO-D-15-0240.1

Hill, C., Ferreira, D., Campin, J.-M., Marshall, J., Abernathey, R., & Barrier, N. (2012). Controlling spurious diapycnal mixing in eddy-resolving height-coordinate ocean models – Insights from virtual deliberate tracer release experiments. *Ocean Modelling*, *45-46*, 14-26. doi: 10.1016/j.ocemod.2011.12.001

Häfner, D., Jacobsen, R. L., Eden, C., Nuterman, R., & Vested, M. H. (2018). Veros v0.1 – a fast and versatile ocean simulator in pure Python. *Geoscientific Model Development*, *11*(8), 3299–3312. doi: 10.5194/gmd-11-3299-2018

Ilicak, M. (2016). Quantifying spatial distribution of spurious mixing in ocean models. *Ocean Modelling*, *108*, 30-38. doi: 10.1016/j.ocemod.2016.11.002

Ilicak, M., Adcroft, A. J., Griffies, S. M., & Hallberg, R. W. (2012). Spurious dianeutral mixing and the role of momentum closure. *Ocean Modelling*, *45-46*, 37-58. doi: 10.1016/j.ocemod.2011.10.003

Ishimwe, A., Deleersnijder, E., Legat, V., & Lambrechts, J. (2023). A split-explicit second order Runge–Kutta method for solving 3D hydrodynamic equations. *Ocean Modelling*, *186*, 102273. doi: 10.1016/j.ocemod.2023.102273

Kar, S. K. (2006). A semi-implicit Runge–Kutta time-difference scheme for the two-dimensional shallow-water equations. *Monthly Weather Review*, *134*(10), 2916-2926. doi: 10.1175/MWR3214.1

Killworth, P. D., Webb, D. J., Stainforth, D., & Paterson, S. M. (1991). The development of a free-surface Bryan–Cox–Semtner ocean model. *J. Phys. Oceanogr.*, *21*, 1333-1348.

Klingbeil, K., Mohammadi-Aragh, M., Gräwe, U., & Burchard, H. (2014). Quantification of spurious dissipation and mixing – Discrete variance decay in a finite-volume framework. *Ocean Modelling*, *81*, 49-64. doi: 10.1016/j.ocemod.2014.06.001

Klöwer, M., Gelbrecht, M., Hotta, D., Willmert, J., Silvestri, S., Wagner, G. L., ... Hill, C. (2024). SpeedyWeather.jl: Reinventing atmospheric general circulation models towards interactivity and extensibility. *Journal of Open Source Software*, *9*(98), 6323. doi: 10.21105/joss.06323

Korn, P. (2017). Formulation of an unstructured grid model for global ocean dynamics. *Journal of Computational Physics*, *339*, 525–552. doi: 10.1016/j.jcp.2017.03.009

Lan, R., Ju, L., Wang, Z., Gunzburger, M., & Jones, P. (2022). High-order multirate explicit time-stepping schemes for the baroclinic-barotropic split dynamics in primitive equations. *Journal of Computational Physics*, *457*, 111050. doi: 10.1016/j.jcp.2022.111050

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., ... Samson, G. (2017). NEMO ocean engine. *Scientific Notes of Climate Modelling Center*, *27*(1288). (Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace no. 27) doi: 10.5281/zenodo.1472492

Marchesiello, P., Debreu, L., & Couvelard, X. (2009). Spurious diapycnal mixing in terrain-following coordinate models: The problem and a solution. *Ocean Modelling*, *26*(3), 156-169. doi: 10.1016/j.ocemod.2008.09.004

Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research: Oceans*, *102*(C3), 5753-5766. doi: 10.1029/96JC02775

NEMO System Team. (2024). *NEMO ocean engine version 5.0: Release notes.* `https://sites.nemo-ocean.io/user-guide/changes.html`. (Accessed: 2024)

Ringler, T., Petersen, M., Higdon, R. L., Jacobsen, D., Jones, P. W., & Maltrud, M. (2013). A multi-resolution approach to global ocean modeling. *Ocean Modelling*, *69*, 211–232. doi: 10.1016/j.ocemod.2013.04.010

Rokhzadi, A., Mohammadian, A., & Charron, M. (2018). An optimally stable and accurate second-order SSP Runge–Kutta IMEX scheme for atmospheric applications. *Journal of Advances in Modeling Earth Systems*, *10*(1), 18-42. doi: 10.1002/2017MS001065

Schifano, N., Vic, C., Gula, J., Molemaker, M. J., & McWilliams, J. C. (2025). Diapycnal mixing and tracer dispersion in a terrain-following coordinate model. *Journal of Advances in Modeling Earth Systems*, *17*(8), e2024MS004768. (e2024MS004768 2024MS004768) doi: 10.1029/2024MS004768

Schlichting, D., Hetland, R. D., & Jones, S. C. (2024). Numerical mixing suppresses submesoscale baroclinic instabilities over sloping bathymetry. *Journal of Advances in Modeling Earth Systems*, *16*(12), e2024MS004321. doi: 10.1029/2024MS004321

Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, *9*(4), 347-404. doi: 10.1016/j.ocemod.2004.08.002

Silvestri, S. (2025). *simone-silvestri/timesteppertestcases.jl: Paper release.* Zenodo. doi: 10.5281/zenodo.17806891

Silvestri, S., Wagner, G., Campin, J.-M., Constantinou, N., Hill, C., Souza, A., & Ferrari, R. (2024). A new WENO-based momentum advection scheme for simulations of ocean mesoscale turbulence. *Journal of Advances in Modeling Earth Systems*, *16*(7), e2023MS004130. doi: 10.1029/2023MS004130

Silvestri, S., Wagner, G. L., Constantinou, N. C., Hill, C. N., Campin, J.-M., Souza, A. N., ... Ferrari, R. (2025). A GPU-based ocean dynamical core for routine mesoscale-resolving climate simulations. *Journal of Advances in Modeling Earth Systems*, *17*(4), e2024MS004465.

(e2024MS004465 2024MS004465) doi: 10.1029/2024MS004465

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L. D., Park, S.-H., & Ringler, T. D. (2012). A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering. *Monthly Weather Review*, *140*(9), 3090–3105. doi: 10.1175/MWR-D-11-00215.1

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., … Powers, J. G. (2008). *A description of the advanced research WRF version 3* (Tech. Rep. Nos. NCAR/TN-475+STR). NCAR. doi: 10.5065/D68S4MVH

Wagner, G. L., Hillier, A., Constantinou, N., Silvestri, S., Souza, A., Burns, K., … Ferrari, R. (2025). Formulation and calibration of CATKE, a one-equation parameterization for microscale ocean mixing. *Journal of Advances in Modeling Earth Systems*, *17*(4), e2024MS004522. (e2024MS004522 2024MS004522) doi: 10.1029/2024MS004522

Wagner, G. L., Silvestri, S., Constantinou, N. C., Ramadhan, A., Campin, J.-M., Hill, C., … Ferrari, R. (2025). High-level, high-resolution ocean modeling at all scales with Oceananigans. *arXiv preprint*. doi: 10.48550/arXiv.2502.14148

Wang, P., Jiang, J., Lin, P., Ding, M., Wei, J., Zhang, F., … Liu, H. (2021). The GPU version of LASG/IAP Climate System Ocean Model version 3 (LICOM3) under the heterogeneous-compute interface for portability (HIP) framework and its large-scale application. *Geoscientific Model Development*, *14*(5), 2781–2799. doi: 10.5194/gmd-14-2781-2021

Wicker, L. J., & Skamarock, W. C. (2002). Time-splitting methods for elastic models using forward time schemes. *Monthly Weather Review*, *130*(8), 2088-2097. doi: 10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2

Yatunin, D., Byrne, S., Kawczynski, C., Kandala, S., Bozzola, G., Sridhar, A., … Schneider, T. (2025). The Climate Modeling Alliance atmosphere dynamical core: Concepts, numerics, and scaling. *ESS Open Archive*. doi: 10.22541/essoar.173246716.58498053/v1