



**HAL**  
open science

# On the Stability and Accuracy of the Harmonic and Biharmonic Isonutral Mixing Operators in Ocean Models

Florian Lemarié, Laurent Debreu, Alexander Shchepetkin, J.C. McWilliams

► **To cite this version:**

Florian Lemarié, Laurent Debreu, Alexander Shchepetkin, J.C. McWilliams. On the Stability and Accuracy of the Harmonic and Biharmonic Isonutral Mixing Operators in Ocean Models. 2012. hal-00665826v1

**HAL Id: hal-00665826**

**<https://hal.science/hal-00665826v1>**

Preprint submitted on 2 Feb 2012 (v1), last revised 26 Apr 2012 (v2)

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

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

# On the Stability and Accuracy of the Harmonic and Biharmonic Adiabatic Mixing Operators in Ocean Models

F. Lemarié<sup>a</sup> L. Debreu<sup>b</sup> A.F. Shchepetkin<sup>a</sup> J.C. McWilliams<sup>a</sup>

<sup>a</sup>*Institute of Geophysics and Planetary Physics, University of California, 405 Hilgard Avenue, Los Angeles, CA 90095-1567, United States*

<sup>b</sup>*INRIA, MOISE Project-Team and Laboratoire Jean Kuntzmann, 51 rue des Mathématiques, 38041 Grenoble Cedex 9, France*

---

## Abstract

Ocean models usually rely on a tracer mixing operator which diffuses along neutral surfaces. This requirement is imposed by the highly adiabatic nature of the oceanic interior, and a numerical simulation needs to respect these small levels of diapycnal mixing to maintain physically realistic results. For non-isopycnal models this is however non-trivial due to the non-alignment of the vertical coordinate isosurfaces with neutral surfaces. This paper considers the numerical solution of initial boundary value problems for harmonic and biharmonic rotated diffusion operators. We provide stability criteria associated with the conventional space-time discretizations of the isoneutral diffusion currently in use in general circulation models. Furthermore, we propose and study possible alternatives to those schemes. A new way to handle the temporal discretization of the rotated biharmonic operator is also introduced. This scheme requires only the resolution of a simple one-dimensional tridiagonal system in the vertical direction to provide the same stability limit than the non-rotated operator. The performance of the various schemes in terms of stability and accuracy is illustrated by idealized numerical experiments of the diffusion of a passive tracer along sloping isopycnals.

*Key words:* Isonneutral Mixing, Diffusion Equations, Time Schemes, Von-Neumann stability analysis, Initial-Boundary Value Problems

---

## 1 Introduction

Most ocean numerical models employ neutral mixing operators either to parameterize the effect of unresolved mesoscale eddies (Gent and McWilliams, 1990; Smith and Gent, 2004), or more basically to control dispersive errors (Lemarié et al., 2012). It is thus very common for non-isopycnic models to implement a rotation of the diffusion tensor in a direction non-aligned with the computational grid. The benefits of a rotated mixing operator in simulating large scale flows are undeniable (e.g., Danabasoglu et al. (1994), Lengaigne et al. (2003)). Much of the improvements brought by the Gent and McWilliams

(1990) parameterization of mesoscale eddies in coarse resolution models are also generally attributed to the adiabatic mixing component of the parameterization which takes the form of a diffusion along isopycnal surfaces (Gent, 2011).

Redi (1982) provided the continuous form of the rotation tensor; however additional efforts were required to proceed to the actual implementation at the discrete level. Several works (Cox, 1987; Danabasoglu and McWilliams, 1995; Griffies et al., 1998; Mathieu et al., 1999; Beckers et al., 1998, 2000) tackled this problem that turned out to be more tedious than expected. The discretization in space raises difficulties to properly conserve the monotonicity (Mathieu and Deleersnijder, 1998; Beckers et al., 2000) and global tracer variance dissipation (Griffies et al., 1998) properties of the continuous operator once the problem is discretized. Moreover, due to the small vertical, relative to horizontal, grid distance typically used in numerical models the vertical and cross terms of the tensor can impose a severe restriction on the time step when explicit-in-time methods are used to advance the rotated operator. This stability problem is alleviated by the use of a standard backward Euler scheme for the vertical component of the tensor (Cox, 1987), at the expense of splitting errors and associated errors in the balance between the active tracer neutral diffusive fluxes (Griffies, 2004). This approach is used in all the state-of-the-art ocean climate models. The existing work on the neutral diffusion has been essentially carried out on the second-order (laplacian) operator and under the small slope approximation (Cox, 1987; Gent and McWilliams, 1990).

Although very few studied so far, a rotated biharmonic may be of interest for high resolution simulations due to its known property of scale selectivity. A biharmonic operator non-aligned with the direction of the computational grid is used in Marchesiello et al. (2009) and Lemarié et al. (2012), and discussed in Griffies (2004). Because global climate models are now targeting increasingly higher horizontal resolution, the question of the viability of such an operator is not only relevant for the regional modeling community but also for the ocean climate community. As an illustration, Hecht (2010) shows that for a  $0.1^{\circ}$  resolution global model the use of a Lax-Wendroff scheme with an implicit diffusion aligned with the horizontal direction leads to too much of a spurious diapycnal mixing in the Equatorial Pacific. This result suggests that even for eddy-resolving simulations an isoneutral mixing operator could be required. This motivates the design of a scale-selective (high-order) rotated operator. This is however not straightforward to maintain the stability of such an operator which can produce undesirable effects like overshooting/undershooting (Delhez and Deleersnijder, 2007) and spurious cabelling processes (Griffies, 2004). To our knowledge, the current implementations of rotated biharmonics are based on an explicit Euler scheme in time with ad-hoc tapering or clipping of the isoneutral slopes to maintain good stability properties (Marchesiello et al., 2009). This approach has however the undesirable effect to allow spurious dianeutral mixing even at places where the slopes are modest and satisfy the small slope approximation.

The aim of this paper is to study a set of space-time discretizations of the rotated harmonic and biharmonic mixing, and to assess them in terms of accuracy, stability and monotonicity violations. One additional constraint we impose to ourselves is accuracy relative to large grid slope ratios (defined as the ratio between the isoneutral slope and the aspect ratio of the computational grid) in order to make the scheme adequate for use in a terrain-following  $\sigma$ -coordinate model. Indeed, problems with  $\sigma$  models are generally more pernicious than with  $z$ -level models because it is not unusual that the slope between the computational grid and the isoneutral direction steepens to be greater than the grid aspect ratio. The paper is organized as follows. In Sec. 2 we introduce the formulation of the isopycnal-mixing problem as well as three different ways to discretize the problem in space. Then Sec. 3 and 4 are respectively dedicated to the temporal discretization of the rotated laplacian and biharmonic operators. Sec. 5 provides the useful details to proceed to the actual implementation of the different schemes in ocean models. Finally, numerical experiments are designed to illustrate the properties of various space-time discretizations in Sec. 6.

## 2 Isopycnal-Mixing Problem Formulation

### 2.1 Continuous Formulation

In this section we briefly introduce the continuous form of the problem under investigation throughout the paper. The three spatial directions are labeled  $x_1, x_2$  for the horizontal coordinates, and  $x_3$  for the vertical coordinate. We note  $q$  the tracer of interest,  $\nabla$  the three-dimensional gradient operator and  $\partial_m = \partial_{x_m}$  ( $m = 1, 3$ ). Following Redi (1982) a mixing operator  $\mathcal{D}_2$  leading to diffusion along isopycnals  $\rho$  with a diffusivity  $\kappa$  can be defined as

$$\mathcal{D}_2 = \nabla \cdot (\mathbf{R} \nabla q) = \sum_{m=1}^3 \partial_m \left( \sum_{n=1}^3 r_{mn} \partial_n q \right), \quad \mathbf{R} = [r_{mn}]_{1 \leq m, n \leq 3}, \quad (2.1)$$

with  $\mathbf{R}$  the following rotation tensor

$$r_{mn} = \kappa \left( \delta_{mn} - \frac{(\partial_m \rho)(\partial_n \rho)}{\|\nabla \rho\|^2} \right), \quad (2.2)$$

where  $\delta_{mn}$  is the conventional Kronecker delta. This form of the tensor shows that the matrix  $\mathbf{R}$  is symmetric positive semi-definite ( $r_{mm} \geq 0$  and  $|r_{mn}| \leq \sqrt{r_{mm} r_{nn}}$ ) and thus defines a diffusion tensor. Moreover, by construction, the diffusive flux which takes the form  $\mathbf{F} = (F^{(1)}, F^{(2)}, F^{(3)})$  with  $F^{(m)} = \sum_{n=1}^3 r_{mn} \partial_n q$  satisfies the orthogonality condition  $\mathbf{F} \cdot \boldsymbol{\rho}_\perp = 0$ , with  $\boldsymbol{\rho}_\perp$  the unit vector in the diapycnal direction. In the limit that horizontal density gradients are much more smaller than vertical gradients (i.e.,  $\sqrt{(\partial_1 \rho)^2 + (\partial_2 \rho)^2} / \partial_3 \rho \ll 1$ ), a simpler form of the tensor, preserving the isopycnal form of mixing, can be devised (Cox, 1987; Gent and McWilliams, 1990). This assumption corresponds to the so-called small slope approximation, and reduces the tensor (2.2) to

$$\begin{cases} r_{mn} = \kappa \delta_{mn} & (1 \leq m, n \leq 2) \\ r_{m3} = r_{3m} = -\kappa \partial_m \rho / \partial_3 \rho & (1 \leq m \leq 2) \\ r_{33} = \kappa ((\partial_1 \rho)^2 + (\partial_2 \rho)^2) / (\partial_3 \rho)^2 \end{cases} \quad (2.3)$$

This formulation of the rotated diffusion under the small slope approximation had also been derived early by Solomon (1971) in a two-dimensional ( $x_1, x_3$ ) case. Throughout this paper we consider the small slope approximation and we allow an anisotropy in the diffusivities (i.e.  $\kappa_1 \neq \kappa_2$ ), the corresponding matrix form of the tensor is

$$\mathbf{R} = \begin{pmatrix} \kappa_1 & 0 & \kappa_1 \alpha_1 \\ 0 & \kappa_2 & \kappa_2 \alpha_2 \\ \kappa_1 \alpha_1 & \kappa_2 \alpha_2 & \kappa_1 \alpha_1^2 + \kappa_2 \alpha_2^2 \end{pmatrix} \quad (2.4)$$

with  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, 0) = -\frac{(\partial_1 \rho, \partial_2 \rho, 0)}{\partial_3 \rho}$  the neutral slope vector. It is straightforward to check that this form of the tensor preserves the symmetry as well as the semi-positive definiteness of the full tensor (2.2). For a laplacian diffusion of a tracer field  $q$  in an unbounded domain  $\Omega = \mathbb{R}^3$ , we can cast the corresponding evolution problem over a time interval  $[0, T]$  under a conservative form

$$\begin{cases} \partial_t q = \mathcal{D}_2(q) = \nabla \cdot (\mathbf{R}\nabla q) = \nabla \cdot \mathbf{F} & \text{in } \Omega \times [0, T], \\ q|_{t=0} = q_0(x_1, x_2, x_3) & \text{in } \Omega, \end{cases} \quad (2.5)$$

with  $q_0$  a given initial condition, and

$$\mathbf{F} = \begin{pmatrix} F^{(1)} \\ F^{(2)} \\ F^{(3)} \end{pmatrix} = \begin{pmatrix} \kappa_1 (\partial_1 q + \alpha_1 \partial_3 q) \\ \kappa_2 (\partial_2 q + \alpha_2 \partial_3 q) \\ \alpha_1 F_1 + \alpha_2 F_2 \end{pmatrix}. \quad (2.6)$$

This problem is here temporarily defined on an unbounded spatial domain in order to facilitate the theoretical study in sections 3 and 4. For bounded domains, we provide the necessary boundary conditions in Sec. 5. Using the notations previously defined, the problem for the rotated biharmonic operator reads

$$\begin{cases} \partial_t q = \mathcal{D}_4(q) = -\mathcal{D}_2(\Psi), & \text{with } \Psi = \mathcal{D}_2(q), & \text{in } \Omega \times [0, T], \\ q|_{t=0} = q_0(x_1, x_2, x_3) & & \text{in } \Omega, \end{cases} \quad (2.7)$$

For the biharmonic case, the diffusivities  $\kappa_1$  and  $\kappa_2$  in (2.4) must be replaced by  $\sqrt{B_1}$  and  $\sqrt{B_2}$  with  $B_1$  and  $B_2$  the hyper-diffusivities. Following Griffies (2004) (Chap. 14), we formulate the biharmonic operator as the composition of two rotated laplacian with coefficients  $\sqrt{B_m}$  ( $m = 1, 2$ ) because this form ensures that, at the continuous level, the variance of  $q$  is strictly dissipated. In the remainder of the paper we focus on the initial value problems (2.5) and (2.7) with the tensor of rotation defined in (2.4).

Unless explicitly said differently, we will use throughout the paper the subscript  $m$  to denote the horizontal directions,  $m = (1, 2)$ .

## 2.2 Spatial Discretization of the Rotated Diffusion Operator

### 2.2.1 Semi-Discrete Considerations

The spatial discretization of the isopycnal mixing operator is a difficult problem which have been thoroughly tackled by Cox (1987); Beckers et al. (1998, 2000); Griffies et al. (1998). We, first, briefly introduce the delicacies associated with the implementation of the operator  $\mathcal{D}_2$ . Because the  $x_1$  and  $x_2$  directions are independent when adopting the small slope approximation, we consider only the problem defined in the  $(x_1, x_3)$  plane. As derived in Beckers et al. (2000), the continuous formulation of the rotated laplacian can be formulated as

$$\mathcal{D}_2(q) = \partial_1 \left( \kappa_1 \frac{\mathcal{J}_1(q, \rho)}{\partial_3 \rho} \right) - \partial_3 \left( \kappa_1 \frac{\partial_1 \rho}{\partial_3 \rho} \frac{\mathcal{J}_1(q, \rho)}{\partial_3 \rho} \right)$$

where  $\mathcal{J}_1(q, \rho) = \partial_1 q \partial_3 \rho - \partial_1 \rho \partial_3 q$  is a Jacobian determinant. The main difficulty resides in the evaluation of the  $\mathcal{J}_1(q, \rho) / \partial_3 \rho$  term on a staggered grid. This term has to be evaluated at the cell interfaces referred as to  $u$ -points and  $w$ -points in Fig. 1. Note that, usually, the isoneutral slope  $\partial_1 \rho / \partial_3 \rho$  used in practice is not the result of a discretization scheme only, but additional ad-hoc constraints taking the form of a smoothing, tapering, or clipping procedure are applied; more details concerning this point are given in Sec. 5. Defining the arithmetic average operators  $\overline{\cdot}^{(1)}$  in the horizontal and  $\overline{\cdot}^{(3)}$  in the vertical, and noting that the natural position of the  $\partial_1 q$  and  $\partial_1 \rho$  (resp.  $\partial_3 q$  and  $\partial_3 \rho$ ) terms is at  $u$ -points (resp.  $w$ -points),

several linear discretizations of the Jacobian determinant can be proposed using a semi-discrete view of the problem:

- A simple discrete analog of the Jacobian  $\mathcal{J}_1$  has been implemented in the early versions of the MOM-GFDL<sup>1</sup> model (Cox, 1987; Danabasoglu and McWilliams, 1995), at  $u$ -points and  $w$ -points the discretization reads

$$\begin{cases} \mathcal{J}_1^u(q, \rho) = \partial_3 \rho \left( \partial_1 q - \partial_1 \rho \frac{\partial_3 \bar{q}^{(3,1)}}{\partial_3 \bar{\rho}^{(3,1)}} \right) \\ \mathcal{J}_1^w(q, \rho) = \partial_3 \rho \left( \partial_1 \bar{q}^{(3,1)} - \partial_1 \bar{\rho}^{(3,1)} \frac{\partial_3 q}{\partial_3 \rho} \right) \end{cases} \quad (2.8)$$

where  $q$  and  $\rho$  are subject to a double averaging in the  $x_3$  and  $x_1$  directions before differencing, which makes this approach prone to a computational mode (Griffies et al., 1998). Note that the  $\partial_3 \rho$  in front of the parenthesis has a passive role here because this is formally  $\mathcal{J}_1(q, \rho)/\partial_3 \rho$  that we aim at discretizing.

- An other way to compute the Jacobian is by differencing before averaging, as proposed in Griffies et al. (1998). In this case we have

$$\begin{cases} \mathcal{J}_1^u(q, \rho) = \partial_3 \rho \left( \partial_1 q - \partial_1 \rho \left[ \frac{\partial_3 q}{\partial_3 \rho} \right]^{(1,3)} \right) \\ \mathcal{J}_1^w(q, \rho) = \partial_3 \rho \left( \overline{\partial_1 q}^{(1,3)} - \overline{\partial_1 \rho}^{(1,3)} \left[ \frac{\partial_3 q}{\partial_3 \rho} \right] \right), \end{cases} \quad (2.9)$$

where the term in brackets could be simply replaced by  $\partial q/\partial \rho$ , however we keep  $\partial_3 \rho$  explicitly to clearly identify the isoneutral slopes.

- To obtain more symmetry between the  $u$  and  $w$  points we can define the following quantity at the corners of the grid cells (sometimes referred as to  $\Psi$ -points)

$$\mathcal{J}_1^\Psi(q, \rho) = \partial_3 \rho \left( \overline{\partial_1 q}^{(3)} - \overline{\partial_1 \rho}^{(3)} \left[ \frac{\partial_3 q}{\partial_3 \rho} \right]^{(1)} \right)$$

which gives

$$\mathcal{J}_1^u(q, \rho) = \overline{\mathcal{J}_1^\Psi(q, \rho)}^{(3)}, \quad \text{and} \quad \mathcal{J}_1^w(q, \rho) = \overline{\mathcal{J}_1^\Psi(q, \rho)}^{(1)}.$$

This discretization has the interesting property to define the isoneutral slopes at  $\Psi$ -points only, which makes it more convenient to handle the tapering, clipping, or smoothing procedure. Indeed, for the schemes (2.8) and (2.9) this procedure has to be done twice, at  $u$  and  $w$  points. However this discretization has the major drawback not to reduce to the classical  $(1, -2, 1)$  stencil when the isoneutral slopes vanish, which disqualifies it.

- We can suggest a last approach which would consist in providing more flexibility to the discretization by introducing two sets of weights  $\nu_n$  and  $\mu_n$  in the problem. We define the operator  $\overline{\cdot}^{(\nu)}$  as the weighted average of the four  $w$ -points surrounding a  $u$ -point, and  $\overline{\cdot}^{(\mu)}$  as the weighted average of the four  $u$ -points surrounding a  $w$ -point. A generalization of scheme (2.9) is

<sup>1</sup> Geophysical Fluid Dynamics Laboratory

$$\begin{cases} \mathcal{J}_1^u(q, \rho) = \partial_3 \rho \left( \partial_1 q - \partial_1 \rho \left\{ \left[ \frac{\partial_3 q}{\partial_3 \rho} \right]^{(\nu)} \right\} \right) \\ \mathcal{J}_1^w(q, \rho) = \partial_3 \rho \left( \overline{\partial_1 q}^{(\mu)} - \overline{\partial_1 \rho}^{(\mu)} \left[ \frac{\partial_3 q}{\partial_3 \rho} \right] \right), \end{cases} \quad (2.10)$$

The parameters  $\nu_n$  and  $\mu_n$  are set by requiring additional properties of the discretization scheme. An example is the LINEAR1 scheme of Beckers et al. (2000) which sets the  $\nu_n$  and  $\mu_n$  coefficients to get the most compact stencil in the cross-isopycnal direction to reduce the amount of spurious diapycnal mixing. This scheme reduces to a (1,-2,1) stencil in the diagonal when the angle between the computational grid and the isoneutral direction is  $\pm 45$  degrees; this property is not satisfied by (2.9).

To our knowledge, three different schemes are currently in use in ocean models. The Nucleus for European Modeling of the Ocean (NEMO, Madec (2008)) and Coupled Large-scale Ice Ocean (CLIO, Goosse et al. (2008); Mathieu et al. (1999)) models use the Cox (1987) discretization (2.8). In NEMO, an horizontal two-dimensional laplacian acting to smooth the isopycnal slopes provides the extra diffusion needed to stabilize the scheme (Cox, 1987; Mathieu and Deleersnijder, 1998; Griffies et al., 1998). In the Modular Ocean Model (MOM, Griffies (2010)) and Parallel Ocean Program (POP, Smith et al. (2010)) the discretization (2.9) based on the tracer/density triads formalism (Fig. 1) is used. A triad is defined as an elementary computational stencil of the rotated operator (Griffies et al., 1998). In MOM, an extra vertical smoothing of the  $\partial_3 \rho$  term is used. Finally, even if not explicitly documented, the scheme (2.10) is implemented in the  $\sigma$ -coordinate Regional Oceanic Modeling System (ROMS, Shchepetkin and McWilliams (2005)) and is routinely used so far to rotate along geopotentials the explicit diffusion in the sponge layers near the open boundaries. We describe below the procedure to compute the weighted averages, and we show that this scheme can also be expressed in a tracer triads formalism. In the present study, the MOM/POP discretization is referred as to TRIADS, the NEMO/CLIO discretization as to COX, and the ROMS discretization as to SW-TRIADS (SW stands for switching).

From our experience, besides the instability identified by Griffies et al. (1998), the TRIADS and COX schemes provide very similar results. The differences are greater between the TRIADS and the SW-TRIADS schemes, we thus focus our study on those two schemes only.

### 2.2.2 Discrete Fluxes

We define  $\delta_1 q_{i+\frac{1}{2},j,k} = q_{i+1,j,k} - q_{i,j,k}$  and  $\delta_3 q_{i,j,k+\frac{1}{2}} = q_{i,j,k+1} - q_{i,j,k}$ . The metric terms  $(\Delta x_2)_{i+\frac{1}{2},j}$  and  $(\Delta x_3)_{i+\frac{1}{2},j,k}$  are the horizontal and vertical measures of the corresponding grid-box interfaces and  $(\Delta x_1)_{i+\frac{1}{2},j}$  is the distance between  $q_{i+1,j,k}$  and  $q_{i,j,k}$ . The vertical index  $k$  varies from  $k = 1$  for the first grid cell next to the ocean floor to  $k = N$  at the surface,  $N$  corresponds to the number of vertical levels of the discretization.

The interfacial  $F^{(1)}$  flux discretized using the TRIADS scheme reads

$$F_{i+\frac{1}{2},j,k}^{(1, \text{triads})} = \kappa_1 (\Delta x_2)_{i+\frac{1}{2},j} (\Delta x_3)_{i+\frac{1}{2},j,k} \left[ \frac{\delta_1 q_{i+\frac{1}{2},j,k}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{\delta_1 \rho_{i+\frac{1}{2},j,k}}{(\Delta x_1)_{i+\frac{1}{2},j}} \cdot \frac{1}{4} \left\{ \frac{\delta_3 q_{i,j,k-\frac{1}{2}}}{\delta_3 \rho_{i,j,k-\frac{1}{2}}} + \frac{\delta_3 q_{i+1,j,k+\frac{1}{2}}}{\delta_3 \rho_{i+1,j,k+\frac{1}{2}}} + \frac{\delta_3 q_{i+1,j,k-\frac{1}{2}}}{\delta_3 \rho_{i+1,j,k-\frac{1}{2}}} + \frac{\delta_3 q_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} \right\} \right]. \quad (2.11)$$

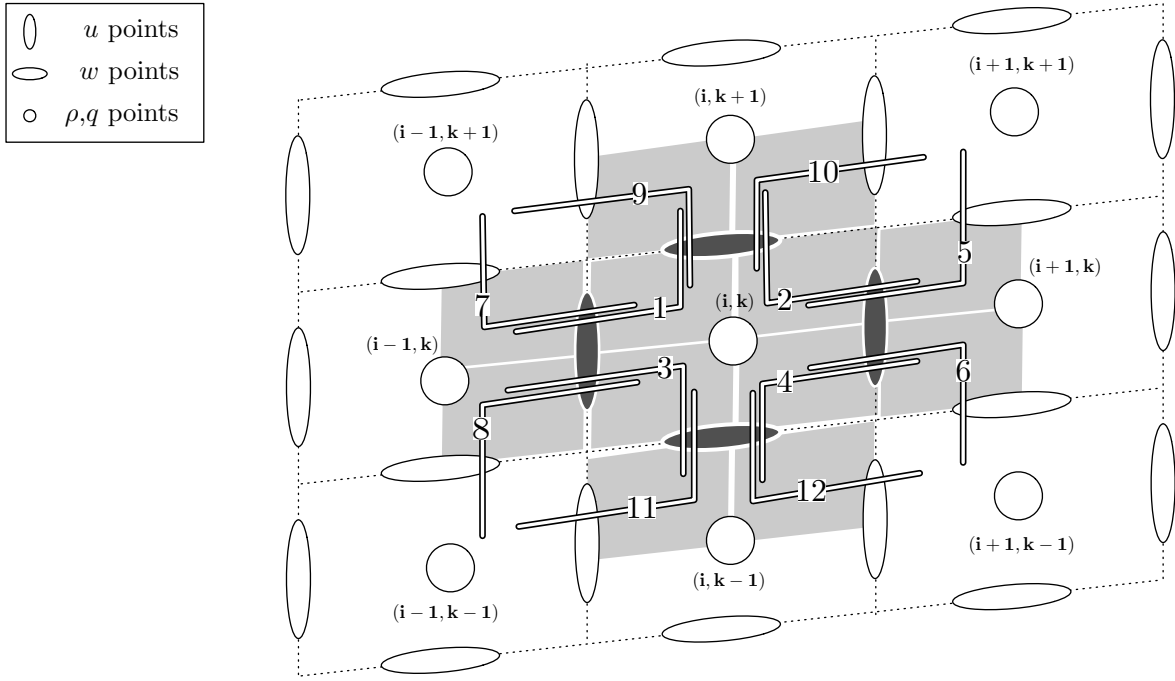


Fig. 1. Grid stencil, constructed upon 12 triads, involved in the computation of the rotated laplacian in the  $(x_1, x_3)$  plane. Each triad has an associated quarter cell (gray shaded areas). The zonal  $F^{(1)}$  and vertical  $F^{(3)}$  components of the neutral diffusive flux are computed at the cell interfaces surrounding the  $(i, k)$  location (dark grey shaded points). The numbering of the triads is meant to be consistent with the one used in Griffies (2004) (Chap. 16).

Using the notations introduced in Fig. 1, we see that the TRIADS scheme uses the four triads labelled 2, 4, 5, and 6 with the same weight  $w = 1/4$ . The derivatives in the  $x_1$ -direction are computed along the horizontal segment of a triad while the derivatives in the  $x_3$ -direction are computed along the vertical segment. Alternatively, the spirit of the LINEAR1 scheme introduced in Beckers et al. (2000) is to weight those triads depending on the orientation of the slope by keeping only two of them. If we assume a stable stratification (i.e.  $\delta_3 \rho_{i,j,k+\frac{1}{2}} < 0, \forall k$ ), the SW-TRIADS scheme reads

$$F_{i+\frac{1}{2},j,k}^{(1,sw-triads)} = \kappa_1 (\Delta x_2)_{i+\frac{1}{2},j} (\Delta x_3)_{i+\frac{1}{2},j,k} \left[ \frac{\delta_1 q_{i+\frac{1}{2},j,k}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{1}{2} \left\{ \frac{w_{i+\frac{1}{2},j,k}^+}{(\Delta x_1)_{i+\frac{1}{2},j}} \left( \frac{\delta_3 q_{i,j,k-\frac{1}{2}}}{\delta_3 \rho_{i,j,k-\frac{1}{2}}} + \frac{\delta_3 q_{i+1,j,k+\frac{1}{2}}}{\delta_3 \rho_{i+1,j,k+\frac{1}{2}}} \right) + \frac{w_{i+\frac{1}{2},j,k}^-}{(\Delta x_1)_{i+\frac{1}{2},j}} \left( \frac{\delta_3 q_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} + \frac{\delta_3 q_{i+1,j,k-\frac{1}{2}}}{\delta_3 \rho_{i+1,j,k-\frac{1}{2}}} \right) \right\} \right]. \quad (2.12)$$

with

$$w_{i+\frac{1}{2},j,k}^+ = \max(\delta_1 \rho_{i+\frac{1}{2},j,k}, 0), \quad w_{i+\frac{1}{2},j,k}^- = \min(\delta_1 \rho_{i+\frac{1}{2},j,k}, 0). \quad (2.13)$$

Depending on the orientation of the slope, either triads 4 and 5 are selected, or triads 2 and 6. This results in a compact four point stencil for the discretization of the  $F_{i+\frac{1}{2},j,k}^{(1,sw-triads)}$  interfacial flux. The cancellation of the contribution of cross-isopycnal points with the SW-TRIADS scheme tends to reduce the amount of spurious diapycnal mixing associated with discretization errors. This point will be exemplified in Sec. 6. The discretization of the vertical flux  $F^{(3)}$  does not raise any additional difficulty for the TRIADS scheme and can be found in Griffies et al. (1998). We show in the next subsection a way to construct the vertical



flux  $F_{i,j,k+\frac{1}{2}}^{(3,\text{sw-triads})}$  so that the SW-TRIADS scheme satisfies a globally diminishing tracer variance.

### 2.3 Global Tracer Variance Dissipation

As shown in Griffies et al. (1998), the TRIADS scheme has been constructed to ensure that the discretized operator globally satisfies the tracer variance dissipation property of the continuous operator. This scheme is designed on the basis of the variational principle. A weak form of the problem under investigation can be defined through

$$\mathcal{G}[q] = -\frac{1}{2} \int_{\Omega} \nabla q \cdot (\mathbf{R} \nabla q) \, d\Omega \quad (2.14)$$

where  $\mathcal{G}[q]$  is a functional whose Frèchet derivative  $\delta\mathcal{G}[q]/\delta q$  gives the diffusion operator. Due to the symmetric positive semi-definite property of the diffusion tensor  $\mathbf{R}$ ,  $\mathcal{G}$  is negative semi-definite. This property implies that the corresponding operator acts to decrease the total tracer variance, indeed Griffies et al. (1998) showed that  $\partial_t \int_{\Omega} q^2 \, d\Omega = 4\mathcal{G}[q]$ . Unlike the COX scheme, the TRIADS scheme provides a negative semi-definite functional at the discrete level, thus ensuring that the corresponding discretized operator is globally strictly dissipative (Griffies et al., 1998; Smith and Gent, 2004; Griffies, 2004). The reader is referred to Griffies (2004) (Chap. 16) for more details about the foundations for the dissipation functional and its discretization.

Consistent with the notations introduced in Griffies (2004) (Chap. 16) the discretization of the functional  $\mathcal{G}$ , in the  $(x_1, x_3)$  plane, is given by

$$\mathcal{G}^{(x_1-x_3)}[q] = -\frac{1}{2} \sum_{i,k} \sum_{n=1}^{12} A(n)V(n) \left[ \frac{\delta_1 q(n)}{\Delta x_1(n)} + \alpha_1(n) \frac{\delta_3 q(n)}{\Delta x_3(n)} \right]^2 \equiv \sum_{i,k} \left( \sum_{n=1}^{12} L_{i,k}^{(n)} \right), \quad (2.15)$$

where the subscripts  $(i, k)$  run over all the cells of the computational domain. In (2.15),  $A(n)$  is the diffusivity associated with triad  $n$ ,  $V(n)$  is the volume of quarter-cell  $n$ ,  $\Delta x_1(n)$  and  $\Delta x_3(n)$  are respectively the length of the horizontal and vertical segments of triad  $n$ . For a given cell  $(i, k)$ , the contribution to the discrete functional is given by  $\sum_{n=1}^{12} L_{i,k}^{(n)}$  which corresponds to 12 nonpositive components associated with 12 quarter cells, built in such a way that there is a unique volume  $V(n)$  and diffusivity  $A(n)$  for each quarter-cell (represented as light gray shaded areas in Fig. 1). This procedure ensures that  $\mathcal{G}^{(x_1-x_3)}[q]$  is negative semi-definite. For example, the contribution of triad 1 to the discretized diffusion operator is given by

$$\frac{\delta L_{i,k}^{(1)}}{\delta q_{i,k}} = -A(1)V(1) \left( \frac{\delta_1 q(1)}{\Delta x_1(1)} + \alpha_1(1) \frac{\delta_3 q(1)}{\Delta x_3(1)} \right) \left( \underbrace{\frac{1}{\Delta x_1(1)}}_{\text{contribution to } F^{(1)}} - \underbrace{\frac{\alpha_1(1)}{\Delta x_3(1)}}_{\text{contribution to } F^{(3)}} \right) \quad (2.16)$$

which shows that when a triad is used to compute  $F^{(1)}$ , the same triad is used to compute  $F^{(3)}$ . The only difference between the TRIADS and the SW-TRIADS schemes lies in the weighting of the 12 triads, the SW-TRIADS scheme simply cancels the contribution of certain triads depending on the orientation of the

isoneutral slope  $\alpha_1$ , which is equivalent to set  $A(n)$  to zero for those triads. This means that once a triad  $n$  is rejected during the computation of  $F^{(1,\text{sw-triads})}$  it can't be used again to compute  $F^{(3,\text{sw-triads})}$  because the diffusivity  $A(n)$  associated with this triad is zero. If we follow this simple rule, we find that there is a unique way to define the vertical interfacial flux  $F_{i,j,k+\frac{1}{2}}^{(3,\text{sw-triads})}$  :

$$\begin{aligned}
F_{i,j,k+\frac{1}{2}}^{(3,\text{sw-triads})} = \kappa_1 \frac{(\Delta x_3)_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} & \left[ \frac{w_{i-\frac{1}{2},j,k}^-}{(\Delta x_1)_{i-\frac{1}{2},j}} \left( \frac{\delta_1 \rho_{i-\frac{1}{2},j,k}}{(\Delta x_1)_{i-\frac{1}{2},j}} \cdot \frac{\delta_3 q_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} - \frac{\delta_1 q_{i-\frac{1}{2},j,k}}{(\Delta x_1)_{i-\frac{1}{2},j}} \right) \right. \\
& + \frac{w_{i+\frac{1}{2},j,k+1}^-}{(\Delta x_1)_{i+\frac{1}{2},j}} \left( \frac{\delta_1 \rho_{i+\frac{1}{2},j,k+1}}{(\Delta x_1)_{i+\frac{1}{2},j}} \cdot \frac{\delta_3 q_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} - \frac{\delta_1 q_{i+\frac{1}{2},j,k+1}}{(\Delta x_1)_{i+\frac{1}{2},j}} \right) \\
& + \frac{w_{i-\frac{1}{2},j,k+1}^+}{(\Delta x_1)_{i-\frac{1}{2},j}} \left( \frac{\delta_1 \rho_{i-\frac{1}{2},j,k+1}}{(\Delta x_1)_{i-\frac{1}{2},j}} \cdot \frac{\delta_3 q_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} - \frac{\delta_1 q_{i-\frac{1}{2},j,k+1}}{(\Delta x_1)_{i-\frac{1}{2},j}} \right) \\
& \left. + \frac{w_{i+\frac{1}{2},j,k}^+}{(\Delta x_1)_{i+\frac{1}{2},j}} \left( \frac{\delta_1 \rho_{i+\frac{1}{2},j,k}}{(\Delta x_1)_{i+\frac{1}{2},j}} \cdot \frac{\delta_3 q_{i,j,k+\frac{1}{2}}}{\delta_3 \rho_{i,j,k+\frac{1}{2}}} - \frac{\delta_1 q_{i+\frac{1}{2},j,k}}{(\Delta x_1)_{i+\frac{1}{2},j}} \right) \right] \times \frac{1}{W}
\end{aligned} \tag{2.17}$$

where  $W$  corresponds to the number of selected triads, i.e. the number of nonzero  $w^\pm$  terms. For  $W = 2$ , it is easy to check that the rotated operator discretized with the SW-TRIADS scheme is based on exactly six triads and is consistent with the functional discretization (2.15) where the contribution of the six other triads is rejected by setting their associated diffusivities to zero. As illustrated in Fig. 2, we have  $W = 2$  for the common situations encountered in ocean models. There are however some degenerated cases (Fig. 2, e) and f) ) for which  $W \neq 2$ . In those very specific cases, the scheme does not satisfy the functional discretization because some triads are used to compute  $F^{(3)}$  with a different weight than it is used to compute  $F^{(1)}$ . We can show easily that when the stratification has a  $2\Delta x_1$  mode, as studied in Griffies et al. (1998), the SW-TRIADS schemes ensure a global tracer variance reduction.

The SW-TRIADS scheme is a special instance of the TRIADS scheme, it thus makes the implementation of this scheme as straightforward. When spatially variable diffusion coefficients are used, each triad must be weighted by the corresponding coefficient. Usually, those coefficients are computed at the horizontal interfacial  $u$  points. In this case, in (2.17) the constant diffusivity  $\kappa_1$  should be dropped and each coefficient  $w^\pm$  must be multiplied by the corresponding diffusivity. Because each triad crosses exactly one  $u$  point (Fig. 1) there is a unique diffusivity associated with a given triad.

It is worth mentioning that the tracer variance diminishing property should not be confused with the total variation diminishing (TVD) property which is monotonicity preserving. As we show in Sec. 6, the rotated operators discretized with the TRIADS or the SW-TRIADS scheme do not preserve monotonicity (Mathieu and Deleersnijder, 1998; Beckers et al., 2000).

Because the rotated biharmonic corresponds to two successive rotated Laplacian, it does not raise any additional difficulties as long as the spatial discretization is concerned. In the two following sections we consider the time integration of the rotated operators, and we assume that the slope vector  $\alpha$  is spatially constant and that the grid is uniform to make the stability analysis tractable. For a constant slope, the discrete approximation  $\widetilde{\mathcal{D}}_2$  of the rotated operator at the position  $(i, j, k)$  is given by

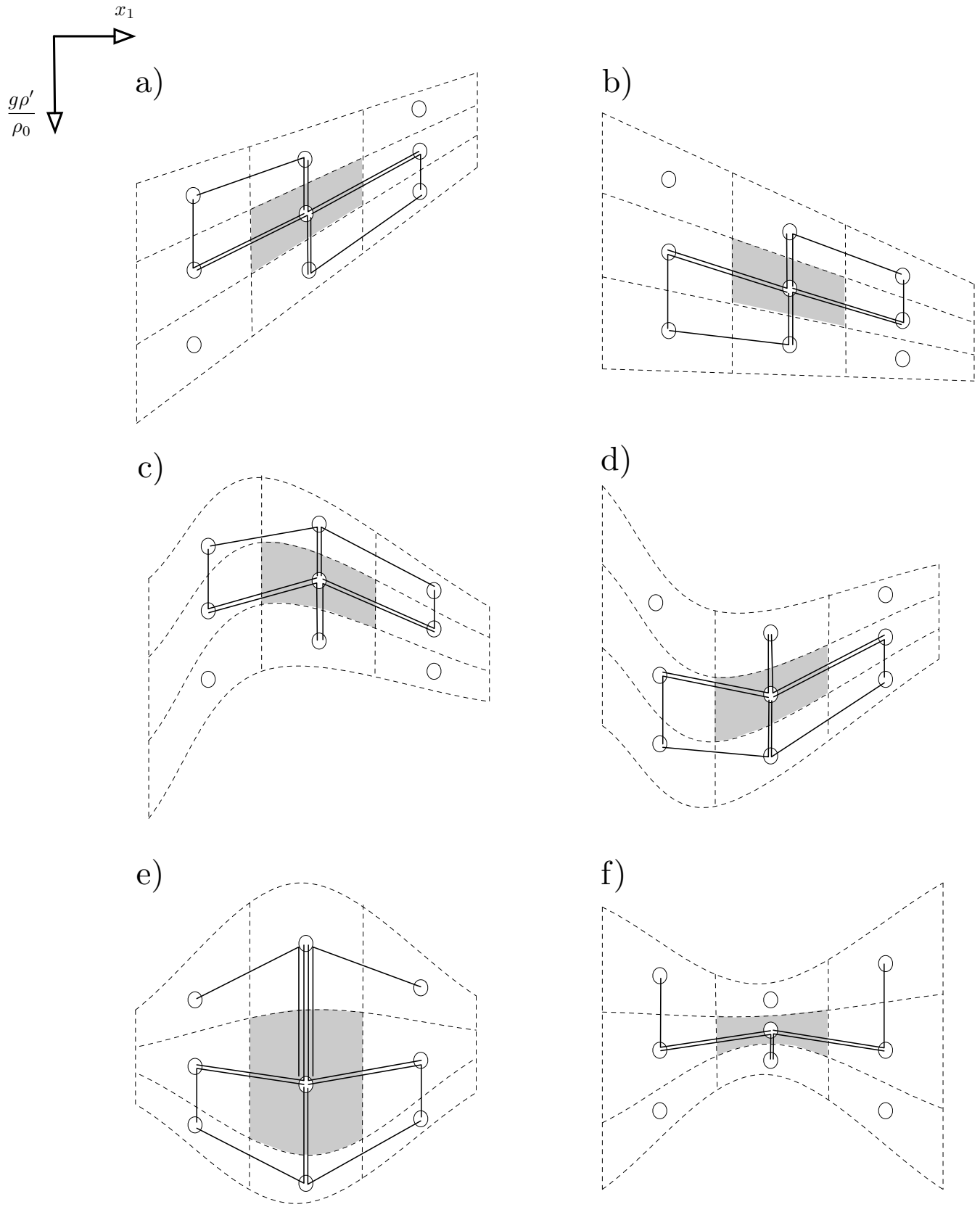


Fig. 2. Computational stencil, represented as triads (black lines), involved in the computation of the rotated laplacian with the SW-TRIADS scheme for different orientations between the computational grid (dashed lines) and the isoneutral direction which corresponds to horizontal lines in the  $(x_1, g\rho'/\rho_0)$  frame. In cases a) to d), 6 triads are selected while 8 triads are selected in e) and 4 in f).

$$\begin{aligned}
(\widetilde{\mathcal{D}}_2)_{i,j,k} &= \frac{F_{i+\frac{1}{2},j,k}^{(1)} - F_{i-\frac{1}{2},j,k}^{(1)}}{\Delta x_1 \Delta x_2 \Delta x_3} + \frac{F_{i,j+\frac{1}{2},k}^{(2)} - F_{i,j-\frac{1}{2},k}^{(2)}}{\Delta x_1 \Delta x_2 \Delta x_3} + \frac{F_{i,j,k+\frac{1}{2}}^{(3)} - F_{i,j,k-\frac{1}{2}}^{(3)}}{\Delta x_3} \\
&= \sum_{p=-1}^1 \sum_{l=-1}^1 \mu_{p,l}^{(1)} q_{i+p,j,k+l} + \sum_{p=-1}^1 \sum_{l=-1}^1 \mu_{p,l}^{(2)} q_{i,j+p,k+l},
\end{aligned} \tag{2.18}$$

with

$$[\mu_{p,l}^{(m)}]_{-1 \leq p,l \leq 1} = \frac{\kappa_m}{\Delta x_m^2} \begin{pmatrix} (\beta_m - 1) \frac{s_m}{2} & s_m(s_m - \beta_m) & (1 + \beta_m) \frac{s_m}{2} \\ 1 - s_m \beta_m & -2(1 + s_m^2) + 2s_m \beta_m & 1 - s_m \beta_m \\ (1 + \beta_m) \frac{s_m}{2} & s_m(s_m - \beta_m) & (\beta_m - 1) \frac{s_m}{2} \end{pmatrix} \tag{2.19}$$

where  $s_m = \frac{\Delta x_m}{\Delta x_3} \alpha_m$  is the grid slope ratio,  $\beta_m \equiv 0$  gives the TRIADS scheme and  $\beta_m = \text{Sign}(1, \alpha_m)$ <sup>2</sup> the SW-TRIADS scheme. The rotated Laplace operator in the  $x_m$ -direction is thus discretized on a centered 9-point stencil if  $\beta_m = 0$  and on a more compact 7-point stencil if  $|\beta_m| = 1$ . As far as the rotated Biharmonic is concerned, the discretization is based on a 25-point stencil with the TRIADS and a 19-point stencil with the SW-TRIADS. We do not investigate this possibility in this study, but the parameter  $\beta_m$  could be used as a degree of freedom to derive alternative properties of the discretization.

The grid slope ratio  $s_m$ , introduced in (2.19), is a key dimensionless parameter that will be used throughout this paper. This parameter corresponds to the ratio between the isoneutral slope  $\alpha_m$  and the aspect ratio  $\Delta x_3/\Delta x_m$ . We thus have  $s_m = \pm 1$  for a  $\pm 45$  degrees slope. Values of  $|s_m|$  greater than 1 generally lead to a degradation of the accuracy of the rotated diffusion due to the need for extrapolation to compute the isoneutral direction. For typical applications with a coarse resolution global climate model,  $|s_m|$  can reach  $\mathcal{O}(10)$  values (e.g.;  $s_m = 10$  for  $\alpha_m = 5 \times 10^{-3}$ ,  $\Delta x_1 = 100$  km, and  $\Delta x_3 = 50$  m). Note that the use of slope clipping or tapering sets an upper bound on  $s_m$  which can't be arbitrarily large in this case. Large grid slope ratios should not be confused with large isoneutral slopes, meaning that large values of  $s_m$  can exist even under the small slope approximation. In the context of a  $\sigma$  to geopotentials rotation, as in Marchesiello et al. (2009),  $s_m$  corresponds to the so-called hydrostatic inconsistency number.

### 3 Time Discretization of the Isonutral Laplacian

#### 3.1 Proposed Schemes

The overall objective of this section is to derive a time-integration scheme whose stability limit is imposed by the horizontal components of the tensor. In this case the constraint on the time step  $\Delta t$  of the temporal discretization would be equivalent between the rotated and the non-rotated operators. Moreover, because we are considering a diffusive process, we do not feel necessary to strive to design a high-order scheme in time, the aim is to keep the study as simple as possible. The rotated laplacian  $\mathcal{D}_2$ , with neutral slopes  $\alpha_m$  and diffusivities  $\kappa_m$ , can be linearly decomposed into a sum of three functions  $\mathcal{D}_2(q) = \sum_{m=1}^3 G_m(q)$ , where

<sup>2</sup> The Sign transfer function is defined as  $\text{Sign}(x, y) = |x|$  if  $y \geq 0$  and  $-|x|$  if  $y < 0$ .

$$\begin{cases} G_m(q) = \partial_m (\kappa_m [\partial_m q + \alpha_m \partial_3 q]) + \partial_3 (\alpha_m \kappa_m \partial_m q) & (m = 1, 2), \\ G_3(q) = \partial_3 \left( [\kappa_1 \alpha_1^2 + \kappa_2 \alpha_2^2] \partial_3 q \right). \end{cases} \quad (3.1)$$

We consider that the piece  $G_0(q) = G_1(q) + G_2(q)$  is always treated explicitly in time-integration schemes, whereas  $G_3$  represent a stiff and unidirectional contribution that can be treated implicitly, if needed. The integration of the  $G_0$  term in an implicit manner would require the solution of a complicated implicit system in the horizontal direction which would be laborious to implement in parallel and would significantly affect the performances of the numerical model. The common practice in climate models is to use a standard backward Euler scheme to advance the vertical  $G_3$  component of the tensor (Cox, 1987). Using the time step  $\Delta t > 0$ , we note  $q^n$  the approximation  $q^n \approx q(t_n)$ , with  $t_n = n\Delta t$ . The semi-discretized version of the scheme introduced by Cox (1987) is

$$q^{n+1} = q^n + \Delta t \{G_0(q^n) + G_3(q^{n+1})\}. \quad (3.2)$$

In the following this scheme will be referred as to (IMP) scheme. Moreover the fully explicit version, i.e. with  $G_3(q^n)$  instead of  $G_3(q^{n+1})$  in (3.2), will be denoted by (EXP).

Following the work of Douglas (1962), Andreev (1967) or Craig and Sneyd (1988) (see also in 't Hout and Welfert (2009) for a review), split schemes such as alternating direction implicit (ADI) have proved valuable in the approximation of the solutions of multi-dimensional parabolic problems with mixed derivatives. This type of scheme is usually implemented with unidirectional implicit corrector steps in each spatial direction to pursue an unconditional stability. Alternatives to this scheme allowing one or more spatial directions to be treated explicitly can be found in Douglas and Gunn (1964); van der Houwen and Verwer (1979) or Hundsdorfer (2002). Those authors propose a multi-stage method : at the first stage, a consistent (explicit) approximation of the operator is evaluated, while all succeeding stages serve to improve the stability. This scheme has been called Method of Stabilizing Corrections (referred as to (MSC) hereafter) in van der Houwen and Verwer (1979) and seems particularly well suited for our problem because it provides a consistent, efficient and easy-to-implement scheme with degrees of freedom to ensure good stability properties. Moreover, we see in Sec. 4 that this approach can be extended to the time-integration of the rotated biharmonic. For the diffusion problem (2.5), the (MSC) scheme reads

$$\begin{cases} q^* = q^n + \Delta t \{G_0(q^n) + G_3(q^n)\} \\ q^{n+1} = q^* + \theta \Delta t \{G_3(q^{n+1}) - G_3(q^n)\} \end{cases} \quad (3.3)$$

where  $\theta > 0$  is a real parameter. Note that for our purpose we only allow a stabilizing correction in the vertical direction. In (3.3),  $\theta = 0$  gives the (EXP) scheme and  $\theta = 1$  the (IMP) scheme. In the next section we study the stability range of the (EXP), (IMP) and (MSC) methods subject to the TRIADS or SW-TRIADS discretizations in space.

### 3.2 Important Results

We give here the important results of our study on the laplacian operator, the associated proof is provided in Sec. 3.3. Those results are given in terms of the parabolic Courant number  $\sigma_m = \kappa_m \Delta t / \Delta x_m^2$ . In numerical models, the value of this parameter usually depends on the time-scheme used to handle the hyperbolic terms. Let us consider a flow-dependent diffusivity  $\kappa_m = \gamma |u_m| \Delta x_m$ , with  $u_m$  the velocity

component in the  $x_m$ -direction and  $\gamma$  a constant; we get  $\sigma_m = \gamma|u|\Delta t/\Delta x_m$ . For example, if the temporal integration is done using a Leap-frog third-order Adams-Moulton (LF-AM3) predictor-corrector method, as in ROMS, the theoretical stability limit is  $|u_1|\Delta t/\Delta x_1 + |u_2|\Delta t/\Delta x_2 \leq 1.587/\pi \approx 0.5$  (Shchepetkin and McWilliams, 2005) which leads to  $\sigma_1 + \sigma_2 \lesssim 0.5\gamma$ . As an illustration, it can be shown that  $\gamma = 1/2$  provides a diffusive term equivalent to the implicit diffusion of the first-order upwind advection scheme, and would correspond to  $\sigma_1 + \sigma_2 \approx 1/4$ .

### *Non-Rotated Laplacian*

When using an Euler forward scheme, the stability limit of the two-dimensional horizontal laplacian, discretized on a five-point stencil<sup>3</sup>, is

$$\sigma_1 + \sigma_2 \leq \frac{1}{2}. \quad (3.4)$$

### *Rotated Laplacian*

- **TRIADS discretization**

- The (EXP) scheme is stable for

$$\sigma_1(1 + s_1^2) + \sigma_2(1 + s_2^2) \leq \frac{1}{2} \quad (3.5)$$

which is always more restrictive than (3.4).

- The stability limit of the (IMP) scheme is given by (3.4).
- The same stability constraint (3.4) is obtained for the (MSC) scheme when  $\theta$  is chosen such that

$$(s_1^2\sigma_1 + s_2^2\sigma_2)\theta = \frac{1}{2} \max \left\{ -1 + 2\Delta t \left[ \sigma_1(1 + s_1^2) + \sigma_2(1 + s_2^2) \right], 0 \right\}. \quad (3.6)$$

The value of  $\theta$  ranges from  $\theta = 0$  when the (EXP) scheme is stable to  $\theta = 1$  when  $\sigma_1 + \sigma_2 = 1/2$ . The amount of implicit diffusion is thus always smaller with the (MSC) scheme compared with the (IMP) scheme.

<sup>3</sup> A nine-point discrete Laplacian operator would provide a stability criterion which is somewhat less restrictive than for the usual five-point one.

- **SW-TRIADS discretization**

- The (EXP) scheme is stable for

$$\sigma_1 \max \{s_1^2, 1\} + \sigma_2 \max \{s_2^2, 1\} \leq \frac{1}{2} \quad (3.7)$$

which shows that for  $s_1^2 \leq 1$  and  $s_2^2 \leq 1$  the stability constraint is the same than the non-rotated operator.

- As for the TRIADS case, the (IMP) scheme is stable if condition (3.4) is satisfied.
- The stability condition (3.4) applies to the (MSC) scheme for

$$\theta = \max \left\{ \frac{|s_1| - 1}{|s_1|}, \frac{|s_2| - 1}{|s_2|}, 0 \right\} \quad (3.8)$$

We now provide the methodology to derive those stability conditions.

### 3.3 Proof Through Linear Stability Analysis

We first assume that the tracer  $q$  can be Fourier decomposed as

$$q(x_1, x_2, x_3, t) = \sum_{\mathbf{k}} A_{\mathbf{k}}(t) \exp \left( i \sum_{m=1}^3 k_m x_m \right), \quad i = \sqrt{-1} \quad (3.9)$$

with  $\mathbf{k} = (k_1, k_2, k_3)$  the three-dimensional wave-vector. Substitution in (2.5) leads to the linear damping equation

$$\frac{dq}{dt} = -\eta q, \quad \text{with } \eta = \sum_{m=1}^2 \kappa_m (k_m + \alpha_m k_3)^2, \quad (3.10)$$

and subsequently

$$q(t_{n+1} = t_n + \Delta t) = \exp(-\eta \Delta t) q(t_n) = \lambda q(t_n). \quad (3.11)$$

$\lambda = e^{-\eta \Delta t}$  provides the exact damping obtained with a “perfect” discretization. We can derive the approximate damping  $\tilde{\lambda}$  provided by the space-time discretization

$$q(t_{n+1}) = \tilde{\lambda}(z_{11}, z_{22}, z_{33}, z_{31}, z_{13}, z_{32}, z_{23}) q(t_n) \quad (3.12)$$

where the  $z_{mn}$  terms are all real and obtained by substitution of discrete Fourier modes in (2.18) and multiplication by  $\Delta t$ :

$$\begin{aligned} z_{mm} &= -2\sigma_m(1 - \cos \phi_m) & m &= 1, 2 \\ z_{33} &= -2(s_1^2 \sigma_1 + s_2^2 \sigma_2)(1 - \cos \phi_3) & & \\ z_{m3} &= z_{3m} = -s_m \sigma_m [\sin \phi_m \sin \phi_3 - \beta_m(1 - \cos \phi_m)(1 - \cos \phi_3)] & m &= 1, 2 \end{aligned} \quad (3.13)$$

with  $\phi_m = k_m \Delta x_m$  ( $|\phi_m| \leq \pi$ ) the normalized Fourier frequencies. Using those notations, we obtain  $\eta \Delta t = \sum_{m=1}^2 \sigma_m (\phi_m + s_m \phi_3)^2$  in (3.11).

The response function of the (MSC) time scheme (3.3) subject to the spatial discretization (2.18-2.19) is

$$\tilde{\lambda} = 1 + \frac{z_0 + z_{33}}{1 - \theta z_{33}}, \quad z_0 = \sum_{m=1}^2 (z_{mm} + z_{m3} + z_{3m}). \quad (3.14)$$

Stability of the scheme under investigation is obtained for  $|\tilde{\lambda}| \leq 1$ .

In the subsequent paragraphs we first show that the condition  $\tilde{\lambda} \leq 1$  is satisfied whatever the value of  $\theta$ , and whatever the spatial discretization. Then, we show that the TRIADS scheme always provides a more restrictive stability range than the SW-TRIADS scheme. Eventually, we derive the requirements to satisfy the constraint  $\tilde{\lambda} \geq -1$  which ensures that the corresponding scheme is stable.

### Upper Bound on $\tilde{\lambda}$

To emphasize the fact that the  $z_{mm}$  terms are negative, we define  $y_m = \sqrt{2\sigma_m(1 - \cos \phi_m)}$ , and  $y_3 = \sqrt{y_{31}^2 + y_{32}^2}$  (with  $y_{3m} = \sqrt{2s_m^2 \sigma_m(1 - \cos \phi_3)}$ ) so that  $z_{mm} = -y_m^2$ . Moreover, we introduce the vector  $\mathbf{v}_m = (\sin \phi_m, \delta_m(1 - \cos \phi_m))$ , with  $\delta_m$  such that  $\beta_m = -\delta_m \delta_3$  and  $|\delta_m| \leq 1$ . We can easily show that

$$z_{m3} = z_{3m} = -s_m \sigma_m (\mathbf{v}_m \cdot \mathbf{v}_3) \quad (3.15)$$

and for  $m = 1, 2$  (thanks to the condition  $|\delta_m| \leq 1$ )

$$\sigma_m \|\mathbf{v}_m\|^2 \leq \sigma_m \left\{ \sin^2 \phi_m + (1 - \cos \phi_m)^2 \right\} = y_m^2, \quad (3.16)$$

where  $\|\bullet\|$  defines the  $\ell_2$ -norm. (3.15) and (3.16) imply that

$$\begin{aligned} -z_{33} - \sum_{m=1}^2 (z_{mm} + z_{m3} + z_{3m}) &= y_1^2 + y_2^2 + y_3^2 + 2s_1 \sigma_1 (\mathbf{v}_1 \cdot \mathbf{v}_3) + 2s_2 \sigma_2 (\mathbf{v}_2 \cdot \mathbf{v}_3) \\ &\geq \sum_{m=1}^2 \sigma_m \|\mathbf{v}_m\|^2 + s_m^2 \sigma_m \|\mathbf{v}_3\|^2 + 2s_m \sigma_m (\mathbf{v}_m \cdot \mathbf{v}_3) \\ &= \sum_{m=1}^2 \sigma_m \|\mathbf{v}_m + s_m \mathbf{v}_3\|^2 \geq 0. \end{aligned} \quad (3.17)$$

This result is sufficient to show that  $\tilde{\lambda} \leq 1$ , indeed thanks to (3.14) we have

$$\tilde{\lambda} = 1 - \left\{ \frac{-z_{33} - \sum_{m=1}^2 (z_{mm} + z_{m3} + z_{3m})}{1 + \theta y_3^2} \right\} \leq 1, \quad (3.18)$$

because we showed that the term in curly brackets is positive. Note that this result is valid whatever  $|\beta_m| \leq 1$  and hence whatever the spatial discretization. This result means that stability of the time discretization is obtained for  $\tilde{\lambda} \geq -1$ . In the following we first show that if this condition is satisfied by the TRIADS scheme the same applies to the SW-TRIADS scheme.



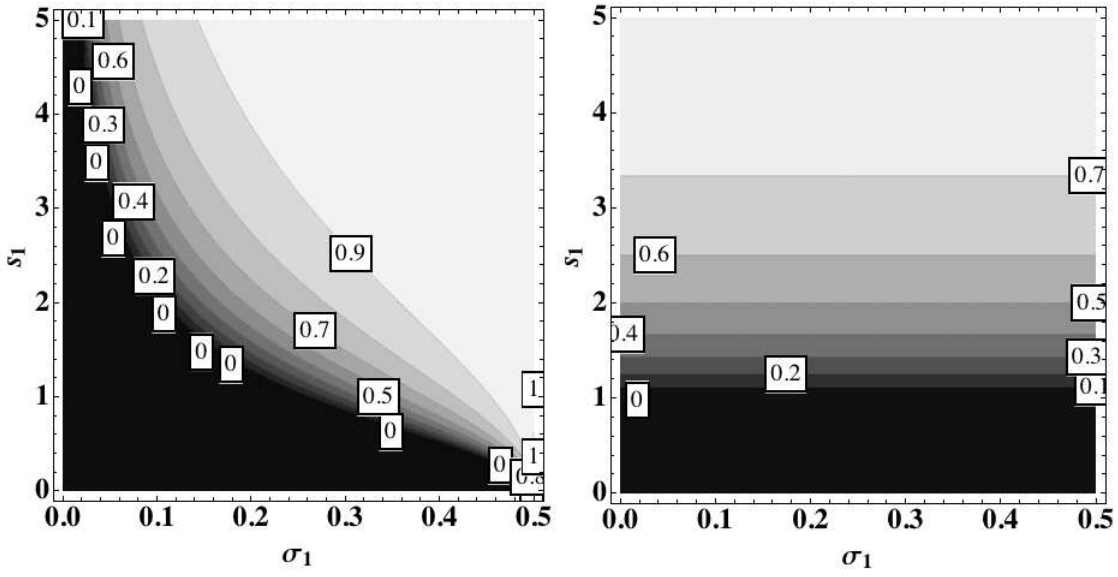


Fig. 3. Optimized values  $\theta^*$  (for  $\sigma_2 = 0$ ) of the stabilizing parameter  $\theta$  obtained with the TRIADS scheme (left) and the SW-TRIADS scheme (right) with respect to  $\sigma_1$  and  $s_1$ .

### Effect of the Cross-Terms

As shown in (3.13), the difference between the two spatial discretizations under consideration appears only in the  $z_{3m}$  terms. The definition of the slope dependent parameter  $\beta_m$  for the SW-TRIADS scheme is such that for all  $s_m$ ,  $\beta_m s_m = |s_m|$ . Substitution in (3.13) leads to

$$-z_{3m} = \sigma_m [s_m \sin \phi_m \sin \phi_3 - |s_m|(1 - \cos \phi_m)(1 - \cos \phi_3)] \leq \sigma_m s_m \sin \phi_m \sin \phi_3. \quad (3.19)$$

This inequality shows that the cross-terms discretized using the SW-TRIADS scheme always provide a less restrictive stability constraint than when using the TRIADS scheme (which corresponds to  $\beta_m = 0$  in (3.13)). It means that the condition  $\tilde{\lambda} \geq -1$  is more difficult to satisfy with the TRIADS discretization. We thus consider in the remainder of this section the most "dangerous" case  $\beta_m = 0$ . We provide in App. B the results obtained for a non-zero value of  $\beta_m$  with the SW-TRIADS discretization.

### Stability Conditions

Starting from (3.14), stability is obtained for

$$\tilde{\lambda} \geq -1 \Leftrightarrow \sum_{m=1}^2 \varphi_m - 2(1 + \theta y_3^2) \leq 0, \quad \varphi_m = y_m^2 + y_{3m}^2 - z_{3m} - z_{m3}. \quad (3.20)$$

Taking  $s_m = 0$  (i.e.  $y_{3m}=z_{3m}=0$ ) and  $\theta = 0$  in (3.20) we find the stability constraint  $y_1^2 + y_2^2 \leq 2$  which corresponds to the usual constraint (3.4) obtained when the horizontal Laplace is integrated using an Euler explicit scheme. Ideally, this is the stability constraint we are targeting for the general case  $s_m \neq 0$  so that the rotated Laplace operator does not involve the use of a smaller time step than the non-rotated Laplace operator. The stability condition (3.20) can be more conveniently written as

$$\varphi_1^\theta + \varphi_2^\theta \leq 2, \quad \text{with } \varphi_m^\theta = \varphi_m - 2\theta y_{m3}^2. \quad (3.21)$$

The stability analysis therefore requires the derivation of a proper upper bound for the  $\varphi_m^\theta$  functions. For clarity, we describe the computation of such an upper bound in App. A, and we recall here only the final result :

$$\varphi_1^\theta + \varphi_2^\theta \leq 2\sigma_1 \mathcal{M}(\theta, s_1) + 2\sigma_2 \mathcal{M}(\theta, s_2) \quad (3.22)$$

where

$$\mathcal{M}(\theta, s_m) = 1 + s_m^2(1 - 2\theta) + \sqrt{(1 + s_m^2)(1 + s_m^2(1 - 2\theta)^2)}. \quad (3.23)$$

Combining (3.21) and (3.22) we get the following criteria for stability,

$$2\sigma_1 \mathcal{M}(\theta, s_1) + 2\sigma_2 \mathcal{M}(\theta, s_2) \leq 2. \quad (3.24)$$

In the fully-explicit (EXP) case (i.e.  $\theta = 0$ ) we get  $\mathcal{M}(0, s_m) = 2(1 + s_m^2)$ , the associated stability criteria reads

$$\sigma_1(1 + s_1^2) + \sigma_2(1 + s_2^2) \leq \frac{1}{2}. \quad (3.25)$$

This result leads to the stability condition (3.5) and is consistent with the result found in Mathieu et al. (1999) (eqn. 31). Furthermore, for  $\theta = 1$  (i.e.; the (IMP) scheme), we get  $\mathcal{M}(1, s_m) = 2$  and thus  $\sigma_1 + \sigma_2 \leq 1/2$  which corresponds to the stability limit of the non-rotated operator. This result shows that by treating implicitly the vertical component of the tensor there is no additional constraint arising from the cross-terms. Finally, for the (MSC) scheme the optimal value of  $\theta$  providing a minimum of implicit diffusion is solution of the equation  $\sigma_1 \mathcal{M}(\theta, s_1) + \sigma_2 \mathcal{M}(\theta, s_2) = 1$ . However, the analytical solution of this equation is extremely complex and would be impractical anyway for real applications. We, therefore, proceed in a more conservative way by noting that for  $\theta \leq 1$

$$\mathcal{M}(\theta, s_m) \leq \mathcal{M}^\dagger(\theta, s_m) = 2(1 + s_m^2) - 2\theta s_m^2. \quad (3.26)$$

It is now straightforward to determine the analytical solution  $\theta^*$  of the equation  $\sigma_1 \mathcal{M}^\dagger(\theta, s_1) + \sigma_2 \mathcal{M}^\dagger(\theta, s_2) = 1$  :

$$(s_1^2 \sigma_1 + s_2^2 \sigma_2) \theta^* = \frac{1}{2} \max \left\{ -1 + 2(1 + s_1^2) \sigma_1 + 2(1 + s_2^2) \sigma_2, 0 \right\}. \quad (3.27)$$

We can check that  $\theta^* = 0$  if the (EXP) scheme is stable and  $\theta^* = 1$  for  $\sigma_1 + \sigma_2 = 1/2$ . In Fig. 3 we show the values of the stabilizing parameter  $\theta^*$  associated with the TRIADS scheme as well as the values of the stabilizing parameter obtained in App. B with the SW-TRIADS. In general, we see that the value of  $\theta$  required to maintain stability of the scheme can be smaller than 1 in numerous cases and hence that splitting errors associated with the (IMP) scheme could be reduced.

### 3.4 Non-oscillatory Scheme

Throughout the previous subsection we have considered the requirement  $|\tilde{\lambda}| \leq 1$  which is a necessary condition for stability. However a more severe condition  $0 \leq \tilde{\lambda} \leq 1$  may be required to ensure a non-oscillatory behavior of the temporal integration (Mathieu et al., 1999). This corresponds to the so-called "no flip-flop" condition, as opposed to the condition  $|\tilde{\lambda}| \leq 1$  which allows a flip-flop behavior in time unlike the exact solution of the problem under investigation (Cushman-Roisin and Beckers, 2011, Chap.

5). Indeed, we see in (3.12) that negative values of  $\tilde{\lambda}$  would allow the solution to oscillate at each time step because  $\tilde{\lambda}, \tilde{\lambda}^2, \tilde{\lambda}^3, \dots$  would change sign. It is straightforward to extend the results described in Sec. 3.3 to the "no flip-flop" case. In the non-rotated case, (3.4) reduces to

$$\sigma_1 + \sigma_2 \leq 1/4, \quad (3.28)$$

which means that the time step must be divided by two. In the rotated case, the (EXP) scheme satisfies the condition  $\tilde{\lambda} \geq 0$  for

$$\sigma_1(1 + s_1^2) + \sigma_2(1 + s_2^2) \leq \frac{1}{4}. \quad (3.29)$$

For the (IMP) scheme, the absence of flip-flop requires

$$\sigma_1 \left(1 + \sqrt{1 + s_1^2}\right) + \sigma_2 \left(1 + \sqrt{1 + s_2^2}\right) \leq \frac{1}{2} \quad (3.30)$$

which is the condition used in the CLIO model to set  $\Delta t$  (Mathieu et al., 1999). The inequality (3.30) shows that the no flip-flop condition is satisfied under a stability constraint more restrictive than (3.28). In the case of the (MSC) scheme, the no flip-flop condition is achieved if (3.28) is satisfied and if  $\theta$  is chosen such that

$$(s_1^2\sigma_1 + s_2^2\sigma_2)\theta^* = \frac{1}{2} \max \left\{ -1 + 4(1 + s_1^2)\sigma_1 + 4(1 + s_2^2)\sigma_2, 0 \right\}. \quad (3.31)$$

For this value of  $\theta$  the scheme satisfies the "no flip-flop" condition if the horizontal terms do, which reduces to (3.28). In (3.31), the value of  $\theta^*$  varies from  $\theta^* = 0$  when (3.29) is satisfied to  $\theta^* = 2$  when  $\sigma_1 + \sigma_2 = 1/4$ . The aim here is not to claim that flip-flops shouldn't be allowed. Indeed, the widely used horizontal laplacian advanced with a forward Euler scheme allows flip-flops for  $1/4 \leq \sigma_1 + \sigma_2 \leq 1/2$  and it didn't turn out to be problematic in practice. This paragraph was meant to illustrate that, unlike the (IMP) scheme which is restricted by the cross-terms, the (MSC) scheme is general enough to ensure additional properties of the time discretization like the absence of flip-flops.

### 3.5 Comments

The literature on multi-dimensional parabolic problems with mixed derivatives has been focused on demonstrating the stability of the method of stabilizing corrections (and/or the ADI method). To our knowledge, there are no systematic studies on the impact of the stabilizing step on the accuracy of the solution. This point is discussed in this section and turns out to be very helpful to anticipate the behavior of the scheme in practical situations.

#### 3.5.1 Spatial Discretization

In Fig. 4 we show the amplification factor obtained when the (EXP) scheme is stable. Because in this case there are no splitting errors, this figure is indicative of the differences between the two spatial discretizations under consideration. To make the interpretation of the results easier, and because we saw earlier that the advective terms are likely to impose the condition  $\sigma_1 + \sigma_2 < 1/4$ , we consider the no flip-flop case. In the figure, we can't formally identify the direction of the computational grid because it represents a wavenumber space, but it is instructive to look at the amplification factor in several directions. Using (3.9), we find that  $\nabla q$  and  $(k_1, k_3)$  are collinear. Moreover it is straightforward to see that the dianeutral direction and  $\rho_{\perp} = (-\alpha_1, 1)$  are collinear too, as well as the isoneutral direction and  $\rho_{\parallel} =$

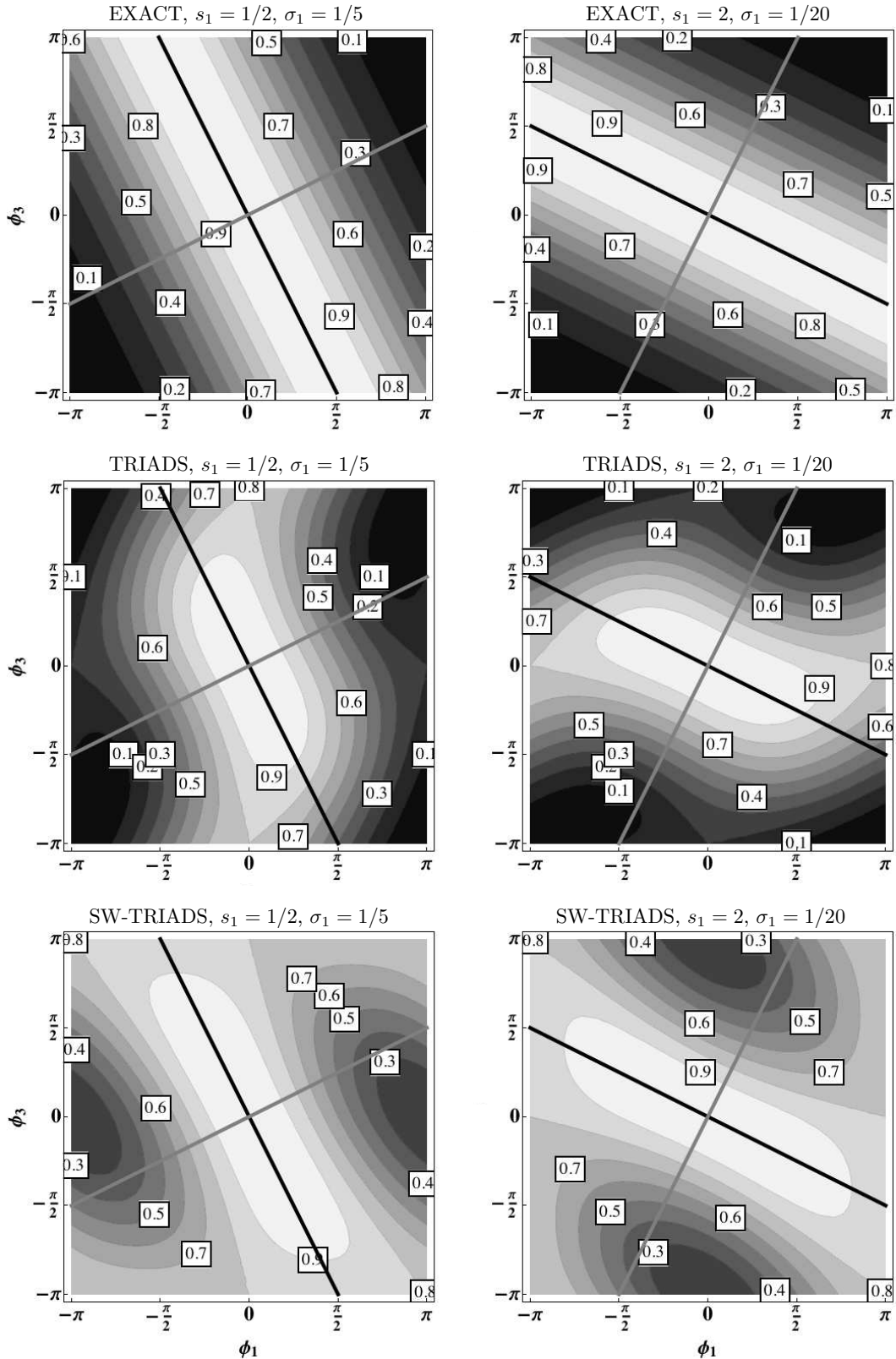


Fig. 4. Exact amplification factor in the  $(\phi_1, \phi_3)$  plane for  $\sigma_1 = 1/4(1 + s_1^2)$  (i.e.; when the (EXP) scheme is stable) with  $s_1 = 1/2$  (top, left) and  $s_1 = 2$  (top, right) in the two-dimensional  $(x_1, x_3)$  case (i.e.  $\kappa_2 = 0$ ). Amplification factor for the (EXP) scheme with the TRIADS discretization (middle, left for  $s_1 = 1/2$  and middle, right for  $s_1 = 2$ ) and with the SW-TRIADS discretization (bottom, left for  $s_1 = 1/2$  and bottom, right for  $s_1 = 2$ ) for the same values of  $\sigma_1$ . The thick black line is  $D_1 = 0$ , the thick gray line is  $D_2 = 0$ .

$(1, \alpha_1)$ . We specifically look at two directions in the  $(\phi_1, \phi_3)$  plane corresponding to two different angles between  $\nabla q$  and the isoneutral/dianeutral direction :

- $D_1 = \phi_1 + s_1\phi_3 = 0$  (i.e.;  $\nabla q \cdot \boldsymbol{\rho}_{\parallel} = 0$ ) corresponds to the direction along which  $\nabla q$  is perpendicular to the isoneutral direction. In this case,  $q$  is constant in the isoneutral direction (because the iso- $q$  lines are perpendicular to  $\nabla q$ ). Because an isoneutral mixing operator should not affect a tracer constant in the neutral direction, we expect the amplification factor  $\lambda$  to be equal to one along this line, otherwise it would indicate that the slope is computed in an inaccurate way and that diapycnal mixing occurs.
- A perpendicular to  $D_1$  is  $D_2 = \phi_3 - s_1\phi_1 = 0$  (i.e.;  $\nabla q \cdot \boldsymbol{\rho}_{\perp} = 0$ ) which corresponds to the direction along which  $\nabla q$  is aligned with the isoneutral direction. In this case, we see from Fig. 4 (top panels) that the exact amplification factor is equivalent to the one of a one-dimensional laplacian, indeed we would have  $\eta\Delta t = \sigma_1(1 + s_1^2)^2\phi_1^2$ .

Along the line  $D_1 = 0$ , the SW-TRIADS always provide a damping in better agreement with the exact one, compared to the TRIADS scheme, thus indicating a more accurate computation of the direction of diffusion for  $s_1 = 1/2$  as well as  $s_1 = 2$ . We see that, for  $\phi_3 = 0$ , all the discretizations have the same behavior and that the mode  $|\phi_1| = \pi$  is effectively damped. This was expected because the difference between the TRIADS and SW-TRIADS schemes lies in the cross-terms which vanish for  $\phi_3 = 0$ . Unlike the SW-TRIADS, the TRIADS scheme has also the property to efficiently damp the checkerboard mode  $(|\phi_1|, |\phi_3|) = (\pi, \pi)$ , whatever the grid slope ratio (Fig. 4). The lack of damping associated with the SW-TRIADS is not problematic to solve the initial value problem (2.5) because we assume some regularity of the initial condition, and the checkerboard mode isn't present in this case. However, when advective terms are considered, they can allow the creation and accumulation of dispersive errors which are expected to be controlled by diffusive processes. We usually do not rely on the rotated operator to damp small scale noise in the vertical direction because the vertical mixing parameterization does it efficiently (otherwise grid-scale noise would arise when an horizontal laplacian is used). In the horizontal direction, those operators are generally the only source of numerical filtering, and we thus expect them to control small-scales noise. For example, for  $s_1 = 1/2$ , the TRIADS scheme does it pretty efficiently while the SW-TRIADS scheme damps the  $2\Delta x_1$  mode ( $|\phi_1| = \pi$ ) only for well-resolved scales in the vertical ( $|\phi_3| \leq \pi/2$ ). Note that it is, however, expected that the  $2\Delta x_3$  mode ( $|\phi_3| = \pi$ ) is already significantly damped by the vertical mixing scheme. The weak damping with the SW-TRIADS scheme can be seen also along the line  $D_2 = 0$ . In this direction, the exact amplification factor and the TRIADS scheme ensure a monotonic damping, indeed  $\lambda$  monotonically decreases when we go from well-resolved to poorly-resolved scales (i.e. when  $|\phi_1|$  or  $|\phi_3|$  increase). This property is not satisfied by the operator discretized with the SW-TRIADS. We thus expect the TRIADS to provide a better control of numerical noise and the SW-TRIADS to provide a more accurate computation of the direction of diffusion. Because spurious diapycnal mixing corresponds to an irreversible mixing error, this type of error can be considered as more harmful than dispersive errors.

### 3.5.2 Space-Time Discretization

We show in Fig. 5 the amplification factor for  $s_1 = 1/2$  and  $s_1 = 2$  when the (MSC) scheme is used with values of  $\theta$  chosen such that there are no flip-flops in time. If we note  $\Delta t_{\rho}$  the time step of the (EXP) scheme and  $\Delta t_0$  the time step of the (MSC) scheme, we get the ratio  $\mu_2$  defined as

$$\mu_2 = \frac{\Delta t_0}{\Delta t_{\rho}} = 1 + s_1^2.$$

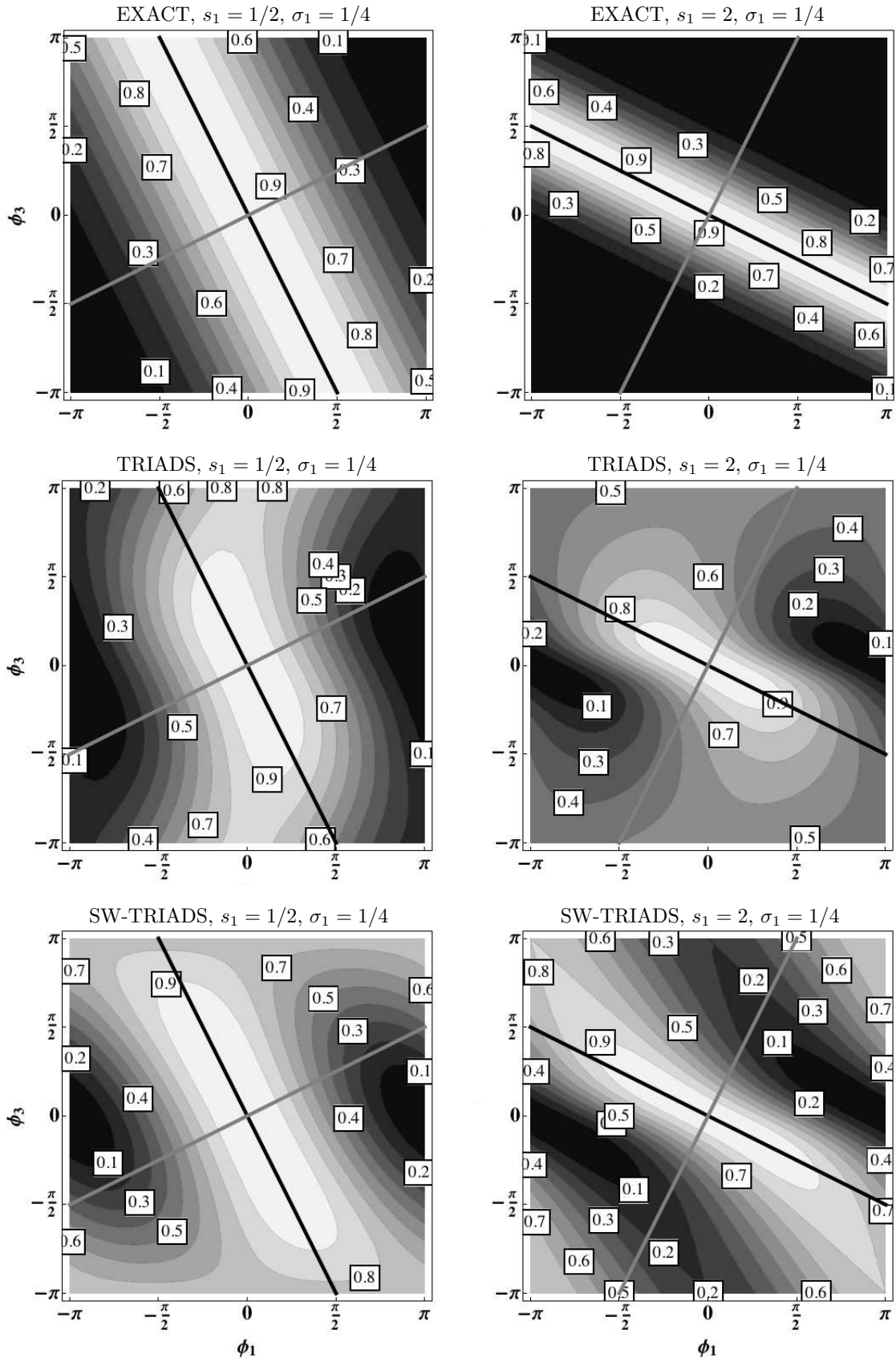


Fig. 5. Same as Fig. 4 for the (MSC) scheme with  $\sigma_1 = 1/4$  (i.e.; when the (EXP) scheme does not satisfy the no flip-flop condition).

For  $s_1 = 1/2$ , we have  $\mu_2 = 5/4$ , and  $\mu_2 = 5$  for  $s_1 = 2$ . When  $\mu$  is close to one, like for  $s_1 = 1/2$ , splitting errors are very small. We see that the left panels in Fig. 5 are very similar to the left panels in Fig. 4. For  $\mu_2 = 5$ , the damping along the line  $D_1 = 0$  increases which indicates that the computation of the direction of diffusion become less and less accurate (Fig. 5, right panels). The space-time discretization is also relatively inaccurate along the line  $D_2 = 0$  and the SW-TRIADS scheme suffers from non-monotonic damping. However, the schemes still perform well as long as the vertical scales are well-resolved (i.e.  $|\phi_3| < \pi/2$ ).

Note that the (MSC) scheme with the TRIADS discretization is particularly interesting either when the absence of oscillations in time is required, or when spatially variable diffusion coefficients are used, otherwise if the diffusivities are constant everywhere and such that  $\sigma_m = 1/2$ , the (MSC) scheme is equivalent to the (IMP) scheme.

## 4 Time Discretization of the Isonutral Biharmonic

### 4.1 Proposed Scheme

As mentioned in the paper's introduction, the isoneutral biharmonic operator has never been thoroughly studied in the literature so far. One possible explanation is that people working on rotated operators were mainly interested in large-scale flows for which the scale-selectivity property is not a priority. Moreover, a very drastic reduction of the time step is required to maintain stability of such an operator. The aim of the present section is to tackle this last point by deriving a time integration scheme enabling an efficient and easy implementation of a rotated biharmonic. To our knowledge, no alternative to the usual explicit Euler scheme has been proposed. It is of course feasible to mimic the (IMP) scheme used for the rotated Laplacian but it would require the solution of a penta-diagonal system, and additional constraints would arise from the cross terms anyway. Indeed, an implicit vertical biharmonic can't stabilize the second order terms in  $x_3$ . Even if it is relatively counter-intuitive we show in the remainder of this section that it is possible to stabilize a biharmonic operator by means of a simple laplacian. An alternative approach could be to use a combination of an implicit laplacian and an implicit biharmonic in the vertical to stabilize separately the second-order and fourth-order terms, but it would increase substantially the complexity of the scheme.

In the same spirit as the method of stabilizing corrections introduced in Sec. 3.1, we thus propose to integrate the rotated biharmonic  $\mathcal{D}_4$  in the following way :

$$\begin{cases} q^* = q^n + \Delta t \mathcal{D}_4(q^n) \\ q^{n+1} = q^* + \Delta t \partial_3 \{ \tilde{\kappa} \partial_3 q^{n+1} - \tilde{\kappa} \partial_3 q^n \}, \end{cases} \quad (4.1)$$

In (4.1), an explicit and consistent approximation of the operator is first computed, followed by a second stage aiming at improving the stability. For the aforementioned reasons, we decided to use a second-order diffusive operator for this second stage. Instead of the stabilizing parameter  $\theta$  used in the previous section, we use here a stabilizing vertical diffusivity  $\tilde{\kappa}$ .

## 4.2 Important Results

We first provide the results and prove them in Sec. 4.3. We keep the notations from the previous section but now with  $\sigma_m$  replaced by  $\sigma_m^{(4)} = \sqrt{\Delta t B_m} / \Delta x_m^2$ , where  $B_m$  is the hyperdiffusivity in the  $x_m$  direction. As in Sec. 3.2 for the laplacian, we can determine an order of magnitude for  $\sigma_m$ . If we assume that  $B_m = \gamma |u_m| \Delta x_m^3$  and an LF-AM3 time-scheme we get  $(\sigma_1^{(4)})^2 + (\sigma_2^{(4)})^2 \lesssim 0.5\gamma$ . The implicit hyperdiffusion associated with a third-order upwind scheme corresponds to  $\gamma = 1/12$  (e.g.; Lemarié et al., 2012) which leads to  $(\sigma_1^{(4)})^2 + (\sigma_2^{(4)})^2 \approx 1/24$ , and  $(\sigma_1^{(4)} + \sigma_2^{(4)})^2 \approx 1/12$  if we assume that  $\sigma_1^{(4)} = \sigma_2^{(4)}$ .

### Non-Rotated Biharmonic

When using an Euler forward scheme, the stability limit of the two-dimensional horizontal biharmonic, is

$$(\sigma_1^{(4)} + \sigma_2^{(4)})^2 \leq \frac{1}{8}. \quad (4.2)$$

This result is also given in Griffies (2004) (Chap. 18).

### Rotated Biharmonic

#### • TRIADS discretization

- The (EXP) scheme is stable for

$$(\sigma_1^{(4)}(1 + s_1^2) + \sigma_2^{(4)}(1 + s_2^2))^2 \leq \frac{1}{8}. \quad (4.3)$$

- The stability constraint of the (MSC) scheme is the same as the stability limit (4.2) for

$$\tilde{\kappa} = 8 \left( \frac{\Delta x_3^2}{\Delta t} \right) (\sigma_1^{(4)} s_1^2 + \sigma_2^{(4)} s_2^2) ((1 + s_1^2)\sigma_1^{(4)} + (1 + s_2^2)\sigma_2^{(4)}) \quad (4.4)$$

#### • SW-TRIADS discretization

- The (EXP) scheme is stable for

$$(\sigma_1^{(4)} \max\{s_1^2, 1\} + \sigma_2^{(4)} \max\{s_2^2, 1\})^2 \leq \frac{1}{8}. \quad (4.5)$$

- The stability constraint of the (MSC) scheme is given by (4.2) for

$$\tilde{\kappa} = 8 \left( \frac{\Delta x_3^2}{\Delta t} \right) (S_1 \sigma_1^{(4)} + S_2 \sigma_2^{(4)}) ((1 + S_1)\sigma_1^{(4)} + (1 + S_2)\sigma_2^{(4)}), \quad (4.6)$$

with  $S_m = \max\{s_m^2 - |s_m|, 0\}$ .

Note that, if we take  $S_m = s_m^2$  in (4.6), we retrieve (4.4). Using the notations introduced in Sec. 2.2, we could write in a generic way  $S_m = s_m(s_m - \beta_m)$ , thus showing that  $S_m$  corresponds to the weight of



the vertical points  $(i, j, k + 1)$  and  $(i, j, k - 1)$  in the stencil (2.19). We now give the proof of the stability results.

### 4.3 Linear Stability Analysis

We, first, define a vertical Courant number  $\tilde{\sigma} = \tilde{\kappa}\Delta t/\Delta x_3^2$  associated with the stabilizing stage in (4.1), and we introduce  $\tilde{y}_3 = \sqrt{2\tilde{\sigma}(1 - \cos \phi_3)}$ . The aim of this stability analysis is to determine a value of  $\tilde{\sigma}$  (and hence  $\tilde{\kappa}$ ) ensuring that the stability constraint of the rotated biharmonic is imposed by the horizontal components of the operator. Note that the exact amplification factor is now given by  $\lambda = \exp(-\eta\Delta t)$  with  $\eta = \sum_{m=1}^2 B_m(k_m + \alpha_m k_3)^4$ , while the amplification factor after discretization is

$$\tilde{\lambda} = 1 - \frac{(y_1^2 + y_2^2 + y_3^2 - z_{13} - z_{31} - z_{23} - z_{32})^2}{1 + \tilde{y}_3^2} = 1 - \frac{(\varphi_1 + \varphi_2)^2}{1 + \tilde{y}_3^2}, \quad (4.7)$$

where  $\varphi_m$  is defined in (3.20). It is straightforward to check that  $\tilde{\lambda} \leq 1$ , stability is thus obtained for  $\tilde{\lambda} \geq -1$ ; this condition is equivalent to

$$(\varphi_1 + \varphi_2)^2 - 2\tilde{y}_3^2 \leq 2. \quad (4.8)$$

We can first easily check that the stability limit of the horizontal biharmonic (i.e., with  $s_m = 0$ ) is given by  $(y_1^2 + y_2^2)^2 \leq 2$ , leading to  $(\sigma_1^{(4)} + \sigma_2^{(4)})^2 \leq 1/8$ . Furthermore, using the upper bound on  $\varphi_m$  found in App. A, we can derive the stability constraint for the forward Euler scheme (i.e.,  $\tilde{\kappa} = 0$ ),

$$(\varphi_1 + \varphi_2)^2 \leq 4 \left( \sigma_1^{(4)} \mathcal{M}(0, s_1) + \sigma_2^{(4)} \mathcal{M}(0, s_2) \right)^2 = 16 \left( \sigma_1^{(4)}(1 + s_1^2) + \sigma_2^{(4)}(1 + s_2^2) \right)^2 \leq 2, \quad (4.9)$$

where  $\mathcal{M}$  is defined in (3.23); (4.9) leads to the expected result (4.3). As an illustration, this stability constraint implies that for  $s_1 = s_2 = 2$ , the time step used with the horizontal biharmonic should be reduced by a factor of 25 when using the rotated biharmonic.

In the general case  $\tilde{\kappa} \neq 0$ , we again use the upper bound on  $\varphi_m$  to get

$$(\varphi_1 + \varphi_2)^2 - 2\tilde{y}_3^2 \leq \mathcal{B}(\theta) = 4 \left( \sigma_1^{(4)} \mathcal{M}(\theta, s_1) + \theta y_{13}^2 + \sigma_2^{(4)} \mathcal{M}(\theta, s_2) + \theta y_{23}^2 \right)^2 - 2\tilde{y}_3^2, \quad (4.10)$$

stability is obtained for  $\mathcal{B}(\theta) \leq 2$ . For simplicity, we take  $\theta = 1$  (i.e.  $\mathcal{M} = 2$ )

$$\begin{aligned} \mathcal{B}(\theta = 1) &= 16 \left( \sigma_1^{(4)} + \sigma_2^{(4)} + (\sigma_1^{(4)} s_1^2 + \sigma_2^{(4)} s_2^2)(1 - \cos \phi_3) \right)^2 - 4\tilde{\sigma}(1 - \cos \phi_3) \\ &= 16 \left( \sigma_1^{(4)} + \sigma_2^{(4)} \right)^2 + (1 - \cos \phi_3) \left\{ 16(\sigma_1^{(4)} s_1^2 + \sigma_2^{(4)} s_2^2) \left[ 2(\sigma_1^{(4)} + \sigma_2^{(4)}) + (\sigma_1^{(4)} s_1^2 + \sigma_2^{(4)} s_2^2)(1 - \cos \phi_3) \right] \right. \\ &\quad \left. - 4\tilde{\sigma}(1 - \cos \phi_3) \right\} \end{aligned} \quad (4.11)$$

The stability constraint  $\mathcal{B}(\theta = 1) \leq 2$  thus reduces to  $(\sigma_1^{(4)} + \sigma_2^{(4)})^2 \leq 1/8$  if the term in curly brackets in (4.11) is negative or zero. This requirement is satisfied for

$$\tilde{\sigma} \geq \sigma^* = 8(\sigma_1^{(4)} s_1^2 + \sigma_2^{(4)} s_2^2) \left( (1 + s_1^2)\sigma_1^{(4)} + (1 + s_2^2)\sigma_2^{(4)} \right), \quad (4.12)$$

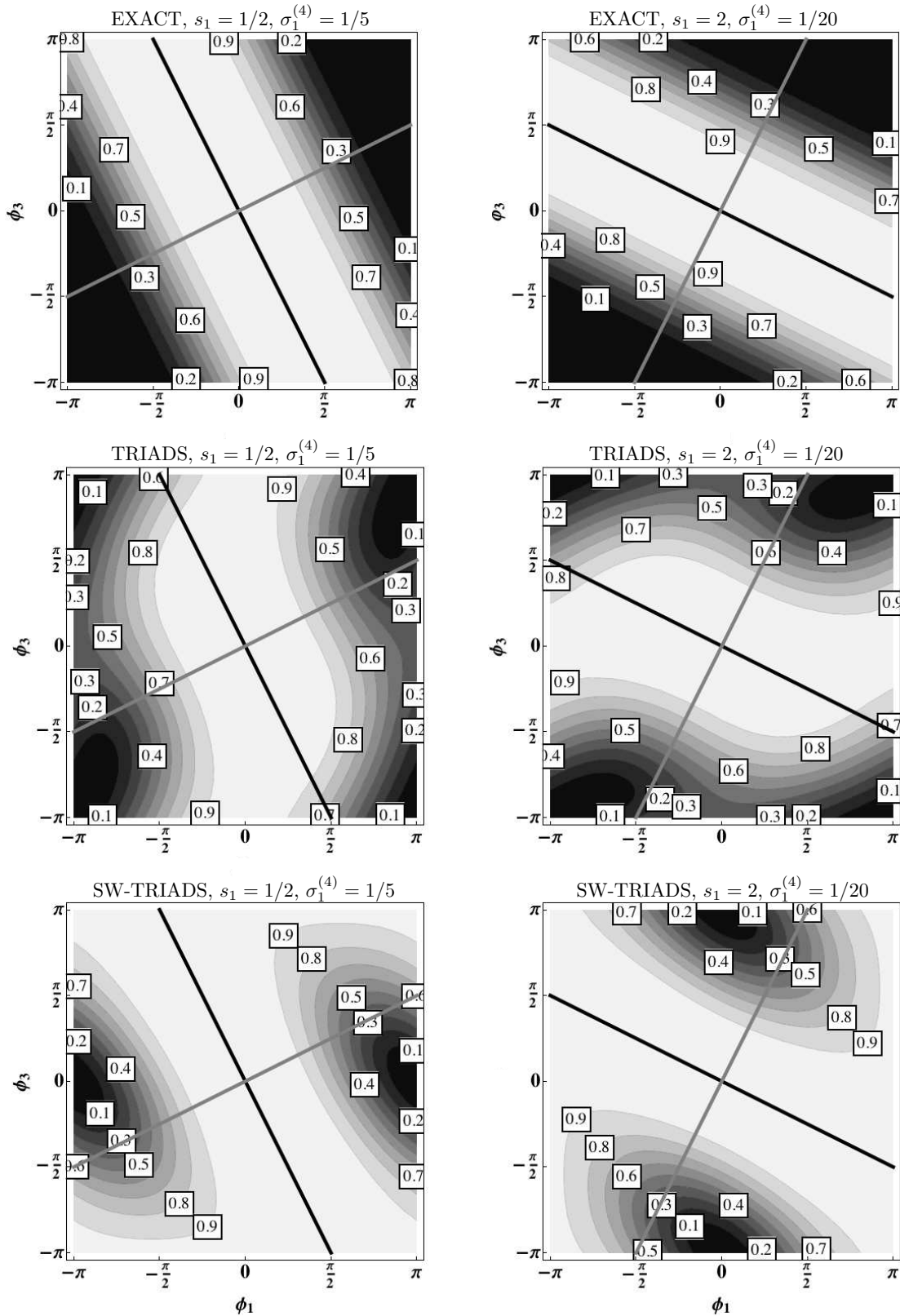


Fig. 6. Exact amplification factor in the  $(\phi_1, \phi_3)$  plane for  $\sigma_1^{(4)} = \sqrt{1/16(1 + s_1^2)^2}$  (i.e.; when the (EXP) scheme is stable) with  $s_1 = 1/2$  (top, left) and  $s_1 = 2$  (top, right) in the two-dimensional  $(x_1, x_3)$  case (i.e.  $\kappa_2 = 0$ ). Amplification factor for the (EXP) scheme with the TRIADS discretization (middle, left for  $s_1 = 1/2$  and middle, right for  $s_1 = 2$ ) and with the SW-TRIADS discretization (bottom, left for  $s_1 = 1/2$  and bottom, right for  $s_1 = 2$ ) for the same values of  $\sigma_1$ . The thick black line is  $D_1 = 0$ , the thick gray line is  $D_2 = 0$ .

which demonstrates (4.4). Less conservative values of  $\tilde{\sigma}$  could be derived (like the one used in Lemarié et al. (2012)), however the algebra becomes very quickly complicated and tedious. The optimal value (4.6) of the stabilizing parameter  $\tilde{\sigma}$  obtained with the SW-TRIADS scheme is provided in App. B.

#### *Non-Oscillatory Scheme*

As for the laplacian, we can derive the necessary constraints to satisfy  $\tilde{\lambda} \geq 0$  instead of  $\tilde{\lambda} \geq -1$ . In this case,  $1/8$  should be replaced by  $1/16$  in (4.2), (4.3), and (4.5). Furthermore, if the stabilizing diffusivity (4.4) is multiplied by two, the (MSC) scheme satisfies the "no flip-flop" condition for  $(\sigma_1^{(4)} + \sigma_2^{(4)})^2 \leq 1/16$ .

#### *4.4 Comments*

We show in Fig. 6 and 7 the amplification factor obtained with a rotated biharmonic for  $s_1 = 1/2$  and  $s_1 = 2$  with the (EXP) and the (MSC) schemes in the no flip-flop case. We can draw the same remarks as for the laplacian: the SW-TRIADS computes more accurately the direction of diffusion, however it suffers from a lack of damping of the checkerboard mode. The ratio  $\mu_4$  between the time step of the (MSC) scheme and the one of the (EXP) scheme is given by

$$\mu_4 = (1 + s_1^2)^2. \quad (4.13)$$

For  $s_1 = 1/2$ , we get  $\mu_4 \approx 1.56$ . In this case, the results are very similar between the (MSC) and the (EXP) schemes. For  $s_1 = 2$ , the ratio becomes stiffer with  $\mu_4 = 25$ . However, the scheme behaves in a similar manner (Fig. 7, right panels) than the (EXP) scheme (Fig. 6, right panels).

#### *4.5 Partial Conclusion and Limitations of Our Approach*

We have shown so far that the laplacian and biharmonic rotated operators can be made stable by combining an explicit-in-time evaluation with a semi-implicit stabilizing step. This approach has the advantage to be extremely easy to implement because it requires only the inversion of a simple tridiagonal system in the vertical. For quite large grid slope ratios ( $s_1 = 2$ ) the scheme performs relatively well but as the slope steepens to  $s_1 = 10$  (Fig. 8) it can suffer from a lack of damping of the smallest resolved scales, especially for the rotated biharmonic. This behavior is reminiscent of the Crank-Nicolson scheme for which the small scale structures are not damped in the limit of large time-steps (Manfredi and Ottaviani, 1999). This problem thus arises when the time step used with the (MSC) scheme is significantly larger than the stability limit of the (EXP) scheme, i.e.; when  $\mu_4$  or  $\mu_2$  are large (for  $s_1 = 10$  we have  $\mu_4 \approx 10000$ ). This lack of damping affects only the vertical direction. The horizontal terms of the tensor, which are advanced explicitly, are still "active". Because the TRIADS scheme requires a larger value of  $\tilde{\kappa}$  to stabilize the time-integration, this lack of damping is generally more pronounced with this scheme. This issue is also more obvious when we try to suppress the flip-flops in time by increasing the value of  $\tilde{\kappa}$  (Fig. 8).

The fact that we restrict ourselves to implicit integrations only in the vertical direction reduces significantly the range of possible methods to try to mitigate this problem. If steep grid slope ratios are expected and accuracy matters, we could combine the (MSC) with a time-splitting approach, not to penalize the all model. When using an Euler explicit scheme, if  $s_1$  reaches  $\mathcal{O}(10)$  values we should decrease the time step by a factor  $10^4$  compared to the time step of the horizontal biharmonic (indeed,  $\mu_4 \approx 10000$  for  $s_1 = 10$ ).

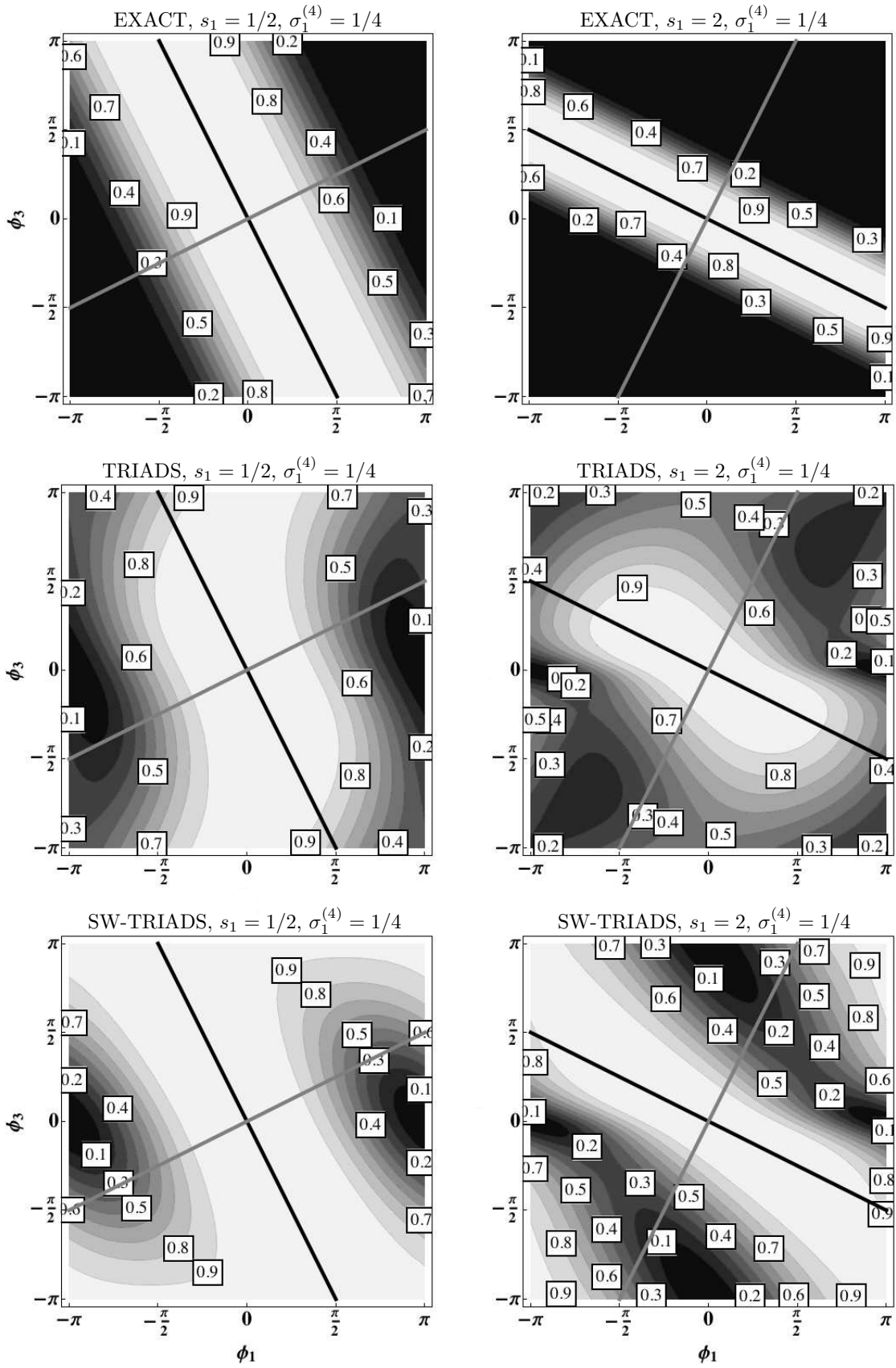


Fig. 7. Same as Fig. 6 for the (MSC) scheme with  $\sigma_1^{(4)} = 1/4$  (i.e.; when the (EXP) scheme does not satisfy the no flip-flop condition).

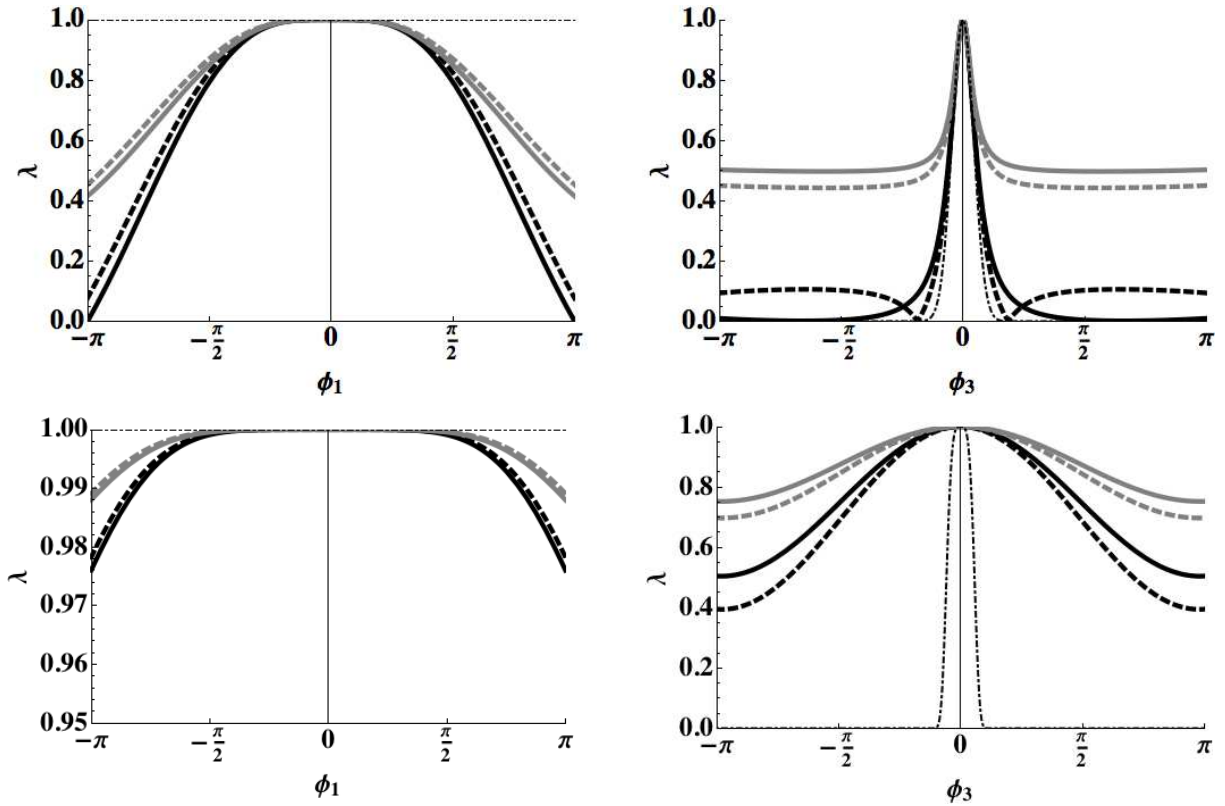


Fig. 8. Amplification factor of the rotated laplacian along the line  $D_1 = 0$  (top, left) and  $D_2 = 0$  (top, right), and of the rotated biharmonic along  $D_1 = 0$  (bottom, left) and  $D_2 = 0$  (bottom, right) for  $s_1 = 10$ ,  $\sigma_1 = 1/4$  and  $\sigma_1^{(4)} = 1/4$ . The gray lines (resp. black lines) are obtained with a stabilizing diffusivity preventing flip-flops (resp. allowing flip-flops) with the (MSC) scheme. The results obtained with the TRIADS (resp. SW-TRIADS) discretization are represented with thick lines (resp. thick dashed lines). The thin black lines correspond to the exact amplification.

Several sub-steps of the (MSC) scheme to advance the rotated operator from time  $n$  to  $n + 1$  would still be significantly more computationally efficient than using an explicit Euler scheme. The use of a small number of sub-step clearly improves the damping properties of the rotated biharmonic which smoothes more efficiently the poorly resolved scales (Fig. 9). Based on the results shown in Fig. 9 we can recommend the use of 4 sub-steps to maintain a good accuracy of the rotated operator for steep grid slope ratios. We checked that for very large values of  $s_1$  ( $s_1 = 10^6$ ), 4 sub-steps are still sufficient to damp the smallest scales (not shown). Note that an even number of sub-steps ensures the absence of flip-flops.

## 5 Implementation of the Proposed Schemes in Existing Models

In this section, we provide additional information required for the implementation of the various schemes introduced in the previous sections.

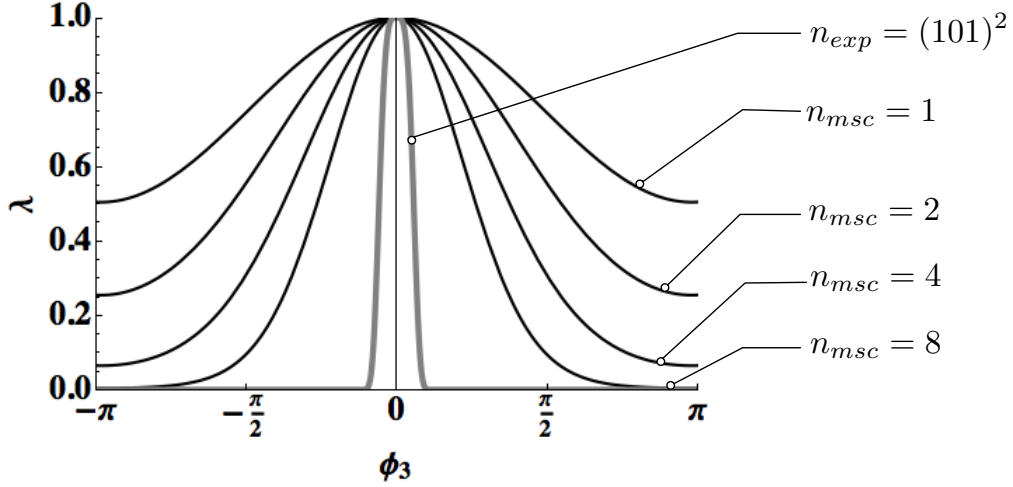


Fig. 9. Amplification factor of the rotated biharmonic along the line  $D_2 = 0$ , for  $s_1 = 10$ . The results are shown for the TRIADS scheme when a sub-stepping of the operator is used with  $n_{msc} = 1, 2, 4$  or  $8$  sub-steps of the (MSC) scheme, and with  $n_{exp} = (1 + s_1^2)^2$  sub-steps of the (EXP) scheme. The case  $n_{msc} = 1$  corresponds to  $\sigma_1^{(4)} = 1/4$ .

### 5.1 Time-Scheme for Diffusive Processes

As mentioned earlier, the majority of the global ocean climate models makes use of a rotated laplacian advanced with the (IMP) scheme. For those models, the implementation of the (MSC) scheme is straightforward and does not require any additional global array for storage. In numerical models like ROMS or MitGCM<sup>4</sup> using a predictor-corrector scheme where provisional values  $q^{n+\frac{1}{2}}$  are given by a predictor step (*e.g.* a Leapfrog-Adams-Moulton interpolation in ROMS or an Adams-Bashforth extrapolation in MitGCM), the (MSC) scheme and the implicit vertical diffusion can be combine with the corrector step in the following way :

$$\begin{cases} q^{n+1,*} = q^n + \Delta t \text{ RHS}^{n+\frac{1}{2}} + \Delta t \mathcal{D}_4(q^n) - \Delta t \partial_3 [\tilde{\kappa} \partial_3 q^n] \\ q^{n+1} = q^{n+1,*} + \Delta t \partial_3 [(K_3 + \tilde{\kappa}) \partial_3 q^{n+1}], \end{cases} \quad (5.1)$$

where  $K_3$  is the vertical eddy-diffusivity given by an appropriate parameterization of the sub-grid scale vertical mixing, and RHS contains all the terms other than diffusive terms. For models using a Leapfrog scheme, like NEMO or POP, the diffusive terms in (5.1) must be evaluated at time  $n - 1$  instead of  $n$ , and  $\Delta t$  must be replaced by  $2\Delta t$  (Madec, 2008; Smith et al., 2010).

### 5.2 Boundary Conditions

We have considered so far a model problem defined on an unbounded domain. For realistic applications, the specification of boundary conditions is required and is based upon two requirements : the mixing operator conserves the tracer content and reduces the global tracer variance. Conservation of the tracer being diffused imposes  $F^{(3)} = 0$  at  $x_3 = \zeta$  (with  $\zeta$  the free-surface) and  $x_3 = H$  (with  $H$  the depth of the water column). To ensure that the boundary conditions are consistent with the global tracer variance

<sup>4</sup> Massachusetts Institute of Technology General Circulation Model, mitgcm.org

diminishing property it is convenient to use the tracer/density triads formalism defined in Sec. 2.3. Using the notations introduced in Fig. 1, we consider that the grid cell  $(i, k)$  is located at the ocean floor. The no-flux boundary condition requires  $F_{i,k-\frac{1}{2}}^{(3)} = 0$  which involves that the triads 3,4,11 and 12 cancel. More generally, each time a given triad crosses the boundary the diffusivity  $A(n)$  associated with this triad in the discrete functional (2.15) should be zero, and this triad can't participate in the computation of the  $F^{(1)}$  interfacial fluxes in the grid-cell next to a boundary. At the bottom (i.e. for  $k = 1$ ) we have

$$F_{i+\frac{1}{2},j,1}^{(1,\text{triads})} = \frac{\kappa_1}{2} (\Delta x_2)_{i+\frac{1}{2},j} (\Delta x_3)_{i+\frac{1}{2},j,k} \cdot \frac{1}{2} \left( \frac{\delta_1 q_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{\delta_1 \rho_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} \frac{\delta_3 q_{i+1,j,\frac{3}{2}}}{\delta_3 \rho_{i+1,j,\frac{3}{2}}} \right. \\ \left. + \frac{\delta_1 q_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{\delta_1 \rho_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} \frac{\delta_3 q_{i,j,\frac{3}{2}}}{\delta_3 \rho_{i,j,\frac{3}{2}}} \right), \quad (5.2)$$

and

$$F_{i+\frac{1}{2},j,1}^{(1,\text{sw-triads})} = \frac{\kappa_1}{2} (\Delta x_2)_{i+\frac{1}{2},j} (\Delta x_3)_{i+\frac{1}{2},j,k} \left( \frac{\delta_1 q_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{w_{i+\frac{1}{2},j,1}^+}{(\Delta x_1)_{i+\frac{1}{2},j}} \frac{\delta_3 q_{i+1,j,\frac{3}{2}}}{\delta_3 \rho_{i+1,j,\frac{3}{2}}} - \frac{w_{i+\frac{1}{2},j,1}^-}{(\Delta x_1)_{i+\frac{1}{2},j}} \frac{\delta_3 q_{i,j,\frac{3}{2}}}{\delta_3 \rho_{i,j,\frac{3}{2}}} \right). \quad (5.3)$$

This approach can be seen as a diffusivity tapering because  $\kappa_1$  is divided by two in the bottom most grid box. An other way to specify a boundary condition is to consider that  $F_{i,j,\frac{1}{2}}^{(3)} = 0$  because  $(\delta_3 \rho_{i,j,\frac{1}{2}})^{-1} = 0$  at the bottom. In this case we get

$$F_{i+\frac{1}{2},j,1}^{(1,\text{triads})} = \kappa_1 (\Delta x_2)_{i+\frac{1}{2},j} (\Delta x_3)_{i+\frac{1}{2},j,k} \left[ \frac{\delta_1 q_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{\delta_1 \rho_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} \cdot \frac{1}{4} \left\{ \frac{\delta_3 q_{i+1,j,\frac{3}{2}}}{\delta_3 \rho_{i+1,j,\frac{3}{2}}} + \frac{\delta_3 q_{i,j,\frac{3}{2}}}{\delta_3 \rho_{i,j,\frac{3}{2}}} \right\} \right] \quad (5.4)$$

and

$$F_{i+\frac{1}{2},j,1}^{(1,\text{sw-triads})} = \kappa_1 (\Delta x_2)_{i+\frac{1}{2},j} (\Delta x_3)_{i+\frac{1}{2},j,k} \left( \frac{\delta_1 q_{i+\frac{1}{2},j,1}}{(\Delta x_1)_{i+\frac{1}{2},j}} - \frac{1}{2} \left\{ \frac{w_{i+\frac{1}{2},j,1}^+}{(\Delta x_1)_{i+\frac{1}{2},j}} \frac{\delta_3 q_{i+1,j,\frac{3}{2}}}{\delta_3 \rho_{i+1,j,\frac{3}{2}}} - \frac{w_{i+\frac{1}{2},j,1}^-}{(\Delta x_1)_{i+\frac{1}{2},j}} \frac{\delta_3 q_{i,j,\frac{3}{2}}}{\delta_3 \rho_{i,j,\frac{3}{2}}} \right\} \right). \quad (5.5)$$

This second approach is recommended only if a specific procedure is used to ensure that  $(\partial_3 \rho)^{-1}$  goes smoothly to zero as a boundary is approached. The same remarks apply at the surface. For realistic applications, the direction of diffusion is generally progressively aligned with the horizontal direction when approaching the surface. This ‘‘boundary rotation’’ is generally done to avoid the rotated operator to interact with the parameterization of sub-grid scale vertical mixing and also to stay consistent with the small slope limit in the well mixed surface layer. If this type of procedure is used, it is safe to consider that  $(\partial_3 \rho)^{-1}$  varies smoothly in the boundary layer and cancels at the surface. We describe in the next section how to handle the transition from isoneutral to horizontal mixing.

In the horizontal, homogeneous Neumann boundary conditions (i.e.,  $\partial_1 q = \partial_1 \rho = 0$ ) are imposed through masking at the coast. Again, because the biharmonic corresponds to two successive laplacian, the specification of the boundary conditions does not imply additional difficulties, and the second boundary condition at the coast required for this operator is simply  $\partial_1^3 q = \partial_1^3 \rho = 0$ . It is straightforward to extend the results of this paragraph to the  $x_2$ -direction. The boundary condition for the stabilizing step in the (MSC)

scheme is an homogeneous Neumann condition to ensure conservation of the tracer content.

### 5.3 Diffusion Slope Limit

A major difficulty when rotating the diffusion in the isoneutral direction is to maintain the numerical integrity of the scheme in poorly stratified regions, including boundary layers and convective regions. Here, the notion of numerical integrity encompasses several aspects : stay consistent with the small slope approximation, smoothly satisfy the boundary condition when conditions (5.4) and (5.5) are used, avoid any infinite slopes and associated infinite fluxes which would introduce numerical instabilities, and avoid any spurious recirculation in the surface layer when an eddy-induced velocity is used. Because the eddy-induced velocity, parameterizing the effect of mesoscale eddies, is usually defined with respect to the vertical derivative of the isoneutral slopes, the type of boundary rotation adopted in the surface layer interacts with the parameterization (Gent and McWilliams, 1990). For this reason, a linear clipping, ensuring a constant eddy-induced velocity in the surface layer, is usually applied to the isoneutral slopes between the base of the surface boundary layer and the surface (Gnanadesikan et al., 2007; Madec, 2008). However, recent advances in the understanding of the role of eddy fluxes in the mixed layer have led to a redesign of the function responsible for the boundary rotation (Ferrari et al., 2008; Fox-Kemper et al., 2008; Colas et al., 2012).

In a more general view, when considering the rotated diffusion independently of the eddy-induced velocity, the important term to keep under control is  $(\partial_3 \rho)^{-1}$ . The use of a specific procedure to control this term to stay consistent with the small slope limit will *de facto* suppress the possibility of infinite slopes. We first define the quantity

$$(\partial_1 \rho)_{i,j,k+\frac{1}{2}}^{\max} = \max \left\{ \frac{|\delta_1 \rho_{i-\frac{1}{2},j,k}|}{(\Delta x_1)_{i-\frac{1}{2},j}}, \frac{|\delta_1 \rho_{i+\frac{1}{2},j,k}|}{(\Delta x_1)_{i+\frac{1}{2},j}}, \frac{|\delta_1 \rho_{i-\frac{1}{2},j,k+1}|}{(\Delta x_1)_{i-\frac{1}{2},j}}, \frac{|\delta_1 \rho_{i+\frac{1}{2},j,k+1}|}{(\Delta x_1)_{i+\frac{1}{2},j}} \right\}$$

and the maximum slope  $\alpha_{\max}$  allowed. The value of  $(\delta_3 \rho)^{-1}$  defined at a cell vertical interface that can be used in practice to compute the isoneutral slopes is given by

$$\left( \delta_3 \rho_{i,j,k+\frac{1}{2}} \right)^{-1} = \max \left\{ -\frac{F(x_3) \cdot \alpha_{\max}}{(\Delta x_3)_{i,j,k+\frac{1}{2}} (\partial_1 \rho)_{i,j,k+\frac{1}{2}}^{\max}}, \min \left\{ -\varepsilon, \left( \delta_3 \rho_{i,j,k+\frac{1}{2}} \right)^{-1} \right\} \right\} \quad (5.6)$$

In the case of an unstable stratification the enforcement of a minimum stratification  $\varepsilon$  rotates the diffusion along the computational grid. In the case of a stable stratification but with large slopes, the slope is bounded by a maximum value  $\alpha_{\max}$ . The procedure (5.6) ensures that the limiting of  $(\delta_3 \rho)^{-1}$  is "felt" the same way when the slopes are computed at  $u$  and  $w$  points. The analytical function  $F$  in (5.6) enables the implementation of a boundary rotation. The value of this function is typically 1 in the oceanic interior and decreases to 0 when approaching the boundaries.

Based on those prescriptions to handle poorly stratified regions and on the ROMS Pacific configuration described in Lemarié et al. (2012), we show in Fig. 10 the maximum value, for each water column, of the quantity  $\mu_4 = \frac{1}{4} ((1 + s_1)^2 + (1 + s_2)^2)^2$  corresponding to the ratio between the time step of the non-rotated biharmonic operator and the time step of the rotated biharmonic operator when the (EXP) scheme is used along with  $B_1 = B_2$ . As mentioned earlier, it is not unusual to get very stiff stability conditions (typically  $\mu_4 > 10000$ ) especially in shallow areas, including coastal regions, and in the Kuroshio extension. The maximum value of  $\mu_4$  is generally found either at the depth of the thermocline (green shaded area in Fig.



10, bottom panel) or at the bottom (blue shaded area in Fig. 10, bottom panel). To generate the figure we used  $\alpha_{\max} = 4/100$  and  $F(x_3) = (x_3/h_{bl})^2 [3 - 2(x_3/h_{bl})]$  for  $x_3 \leq h_{bl}$  ( $F(x_3) = 1$  otherwise) where  $h_{bl}$  is the depth of the surface boundary layer given by the parameterization of the vertical mixing. The maximum value of  $\mu_4$  we show in Fig. 10 is generally not obtained for  $\alpha_m = \alpha_{\max}$  but for places where high vertical resolution coincides with sloping isopycnals. A decrease of the parameter  $\alpha_{\max}$  (e.g. from 4/100 to 1/100) would give very similar values for  $\mu_4$ . Because it is not unusual to get very large values of  $\mu_4$  in realistic configurations, it makes it challenging to maintain a good accuracy of the rotated operator under all circumstances when the time step is set by the stability condition of the non-rotated operator.

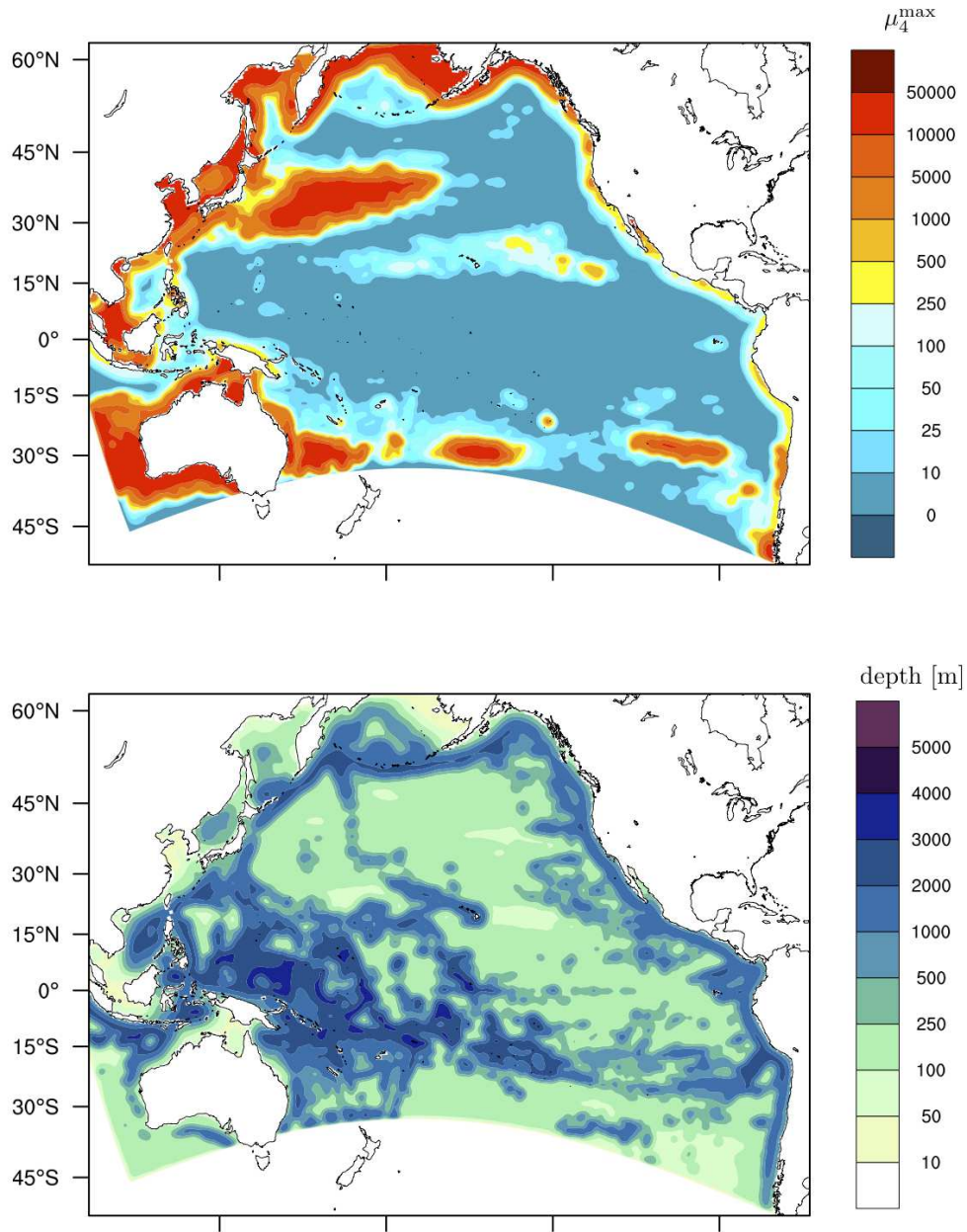


Fig. 10. Annual mean of the maximum value of the ratio  $\mu_4$  between the time step of the non-rotated biharmonic operator and the time step of the rotated biharmonic operator for each water column (top), and depth of the maximum (bottom).  $\mu_4$  is first computed using seasonal averages and is then averaged to get the annual mean.

Resolution	Scheme	Time Step	Maximum slope parameter ( $s_1$ )		$\Delta t_0/\Delta t_\rho$			
			SMALL	LARGE	TRIADS		SW-TRIADS	
					$\mathcal{D}_2$	$\mathcal{D}_4$	$\mathcal{D}_2$	$\mathcal{D}_4$
1024x96	(EXP)	$\Delta t_\rho$	0.08	0.35	1	1	1	1
256x48	(MSC)	$\Delta t_0$	0.16	0.70	1.5	2.25	1	1
128x48	(MSC)	$\Delta t_0$	0.32	1.4	2.96	9	1.96	3.84
64x24	(MSC)	$\Delta t_0$	0.32	1.4	2.96	9	1.96	3.84
32x24	(MSC)	$\Delta t_0$	0.64	2.8	8.8	77	7.84	61.5

Table 1

Description of the numerical experiments discussed in Sec. 6. Those experiments are carried out using the rotated laplacian (with  $\kappa_1 = 5 \text{ m}^2 \text{ s}^{-1}$ ) and the rotated biharmonic (with  $B_1 = 5 \times 10^{-1} \text{ m}^4 \text{ s}^{-1}$ ) with either the SW-TRIADS or the TRIADS spatial discretizations.  $\Delta t_\rho$  is given by the stability limit (3.5) for the laplacian operator and (4.3) for the biharmonic operator.  $\Delta t_0$  corresponds to the time step restriction arising from the horizontal component of the rotated operators; i.e.  $\Delta t_0 = \Delta x_1^2/2\kappa_1$  for the diffusion, resp.  $\Delta t_0 = \Delta x_1^4/8B_1$  for the hyperdiffusion. The ratio  $\Delta t_0/\Delta t_\rho$  is given for the LARGE experiments with the rotated laplacian  $\mathcal{D}_2$  and for the rotated biharmonic  $\mathcal{D}_4$ .

## 6 Numerical Results

In this section, we present a set of idealized experiments to investigate the behavior of the various space-time discretization schemes we introduced earlier in the paper.

### 6.1 Experimental Setup

We have implemented a two-dimensional  $(x_1, x_3)$  testcase defined on the computational domain  $\Omega = [0, 1] \times [0, 1]$  with a time-independent stratification

$$\rho_\xi(x_1, x_3) = -\tanh\left(5 \left\{x_3 - \frac{1}{4} - \xi 8\pi^3 x_1^3 \left[\sin(\pi x_1) - \frac{1}{2} \sin(2\pi x_1)\right]^2\right\}\right). \quad (6.1)$$

The slope associated with this stratification is

$$\alpha_1 = -\frac{\partial_1 \rho_\xi}{\partial_3 \rho_\xi} = \xi \cdot R(x_1), \quad (6.2)$$

where  $R(x) = 64\pi^3 x^2 \cos(\pi x/2) \sin(\pi x/2)^5 (2\pi x + 4\pi x \cos(\pi x) + 3 \sin(\pi x))$ . In (6.1),  $\xi$  is a stiffness parameter used to define the amplitude of the angle between the computational grid and the isopycnal direction. Choosing two values of  $\xi$ , we build two set of experiments which are described in Tab. 1 : a first one with a small grid slope ratio between the computational grid and the isopycnal direction (hereafter SMALL) and a second with a large grid slope ratio (hereafter LARGE). The largest value we consider is  $s \approx 3$  which, for example, would correspond to  $\alpha_1 = 7.5 \times 10^{-4}$ ,  $\Delta x_1 = 100 \text{ km}$  and  $\Delta x_3 = 25 \text{ m}$ .

The stratification (6.1) is shown in Fig. 11 along with the initial condition

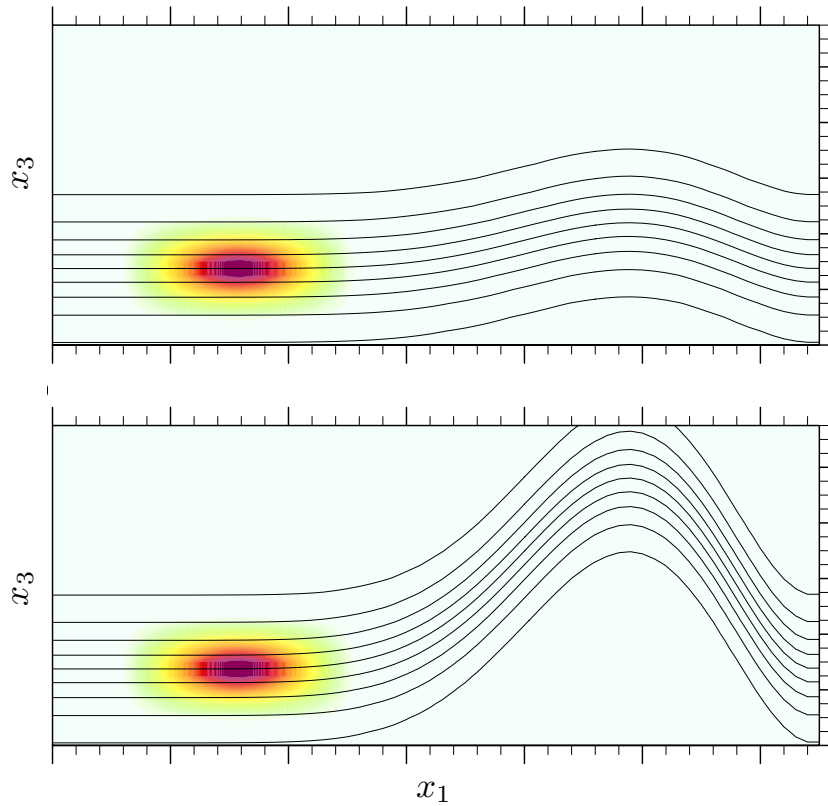


Fig. 11. Two-dimensional  $(x_1, x_3)$  initial conditions for different experiments of a passive tracer (shaded) diffused along a density field (contours) with a moderate slope between the computational grid and the isopycnals (top), and with a large slope (bottom).

$$q_0(x_1, x_3) = \frac{1}{4} \left[ \cos \left( (20 x_3 - 5) \frac{\pi}{3} \right) + 1 \right] \left[ \cos \left( (20 x_1 - 5) \frac{\pi}{3} \right) + 1 \right] \mathbf{I}_{[1/10, 4/10]}(x_1) \mathbf{I}_{[1/10, 4/10]}(x_3), \quad (6.3)$$

where  $\mathbf{I}_\Omega(x)$  is the indicator function<sup>5</sup>. An additional case with a density field with a  $2\Delta x_1$  structure is defined to investigate the global tracer variance diminishing property of the various discretizations under consideration (Fig. 11, top panel). All the simulations described in Tab. 1 are run with the same diffusivity  $\kappa_1 = 5 \text{ m}^2 \text{ s}^{-1}$  and hyperdiffusivity  $B_1 = 5 \times 10^{-1} \text{ m}^4 \text{ s}^{-1}$ .

## 6.2 Accuracy of the Rotated Operators

The accuracy of different space-time discretizations is checked by computing a high resolution solution on a 1024x96 grid using an explicit Euler scheme in time. This solution is taken as a solution of reference for a set of four coarser grids, 256x48, 128x48, 64x24, and 32x24. With the exception of the reference solution, all the simulations are done using the (MSC) scheme with the stabilizing parameters (3.6), (3.8), (4.4) and (4.6) with  $\kappa_2 = 0$  (resp.  $B_2 = 0$ ). To build this hierarchy of grids we have considered that when the horizontal resolution is refined by  $n^2$  the vertical resolution is refined by  $n$  (Tab. 1). This rule is meant to be quite consistent with the usual practice in climate models. Due to the excessive computational cost

<sup>5</sup>  $\mathbf{I}_\Omega(x) = 1$ , if  $x \in \Omega$ ;  $\mathbf{I}_\Omega(x) = 0$  otherwise.

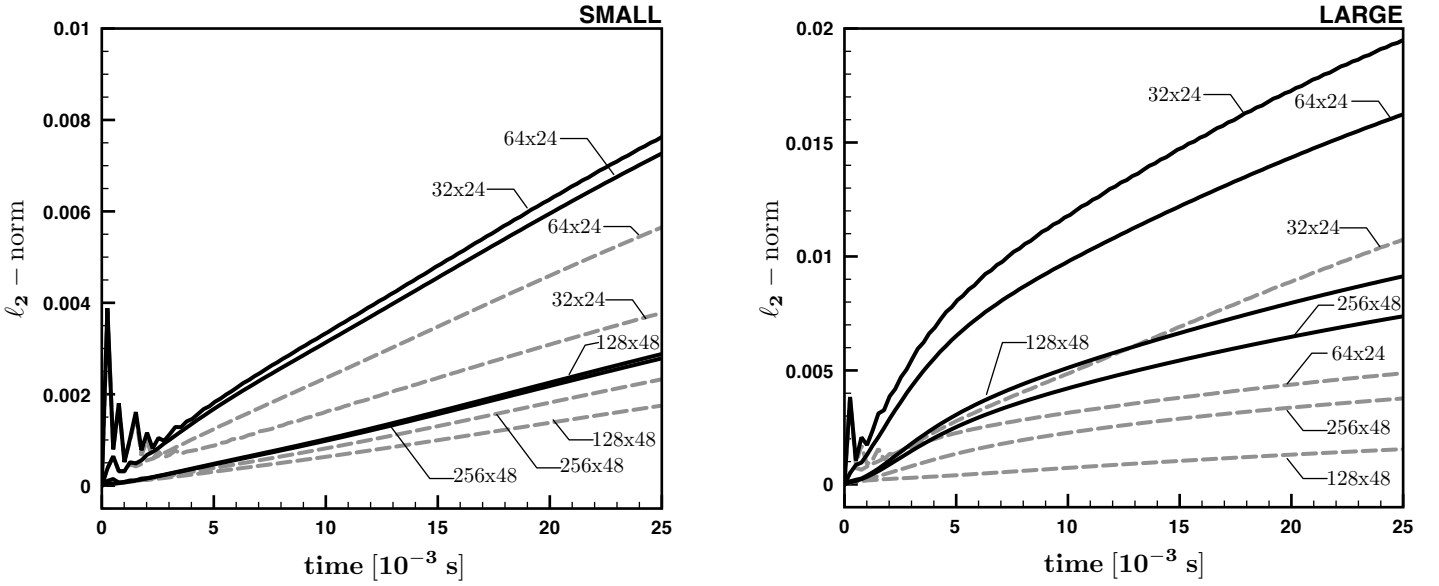


Fig. 12. Time history of the error  $\ell_2$ -norm for different grid resolutions, in the SMALL experiments (left) and in the LARGE experiments (right). The black lines correspond to the results obtained with the TRIADS scheme, and the dashed gray lines to the results obtained with the SW-TRIADS. Note the different vertical axis between the two panels.

of the explicit-in-time rotated biharmonic on the 1024x96 grid, we use the 256x48 grid as a solution of reference in this case.

The stratification (6.1) and the initial condition (6.3) are first defined on the reference 1024x96 grid and then a coarsening procedure is used to provide it to the entire hierarchy of grids. The same coarsening procedure is then applied to the solution of reference to compute the  $\ell_2$ -norm of the error for each grid of the hierarchy. This procedure ensures that the error is initially zero. The results obtained for the SMALL and the LARGE experiments are shown in Fig. 12. When using the TRIADS scheme, the error is monotonically decreasing when the grid is refined. For large grid slope ratios this discretization leads to a significant amount of spurious diapycnal diffusion due to discretization errors (Fig. 13, g) and i). The SW-TRIADS scheme systematically leads to smaller errors whatever the grid resolution and the grid slope ratio (Fig. 12). However, the evolution of the errors with the resolution is relatively uncommon. For example, in the LARGE experiments, the error is smaller on the 128x48 grid than on the 256x48 grid. This is explained by the fact that the errors of discretization associated with the SW-TRIADS scheme do not monotonically increase when the grid slope ratio  $s_1$  increases, as this is the case for the TRIADS scheme. This point will be further discussed in Sec. 6.4 but it can be seen from (2.19) that the discretization reduces to the classical 1,-2,1 stencil for  $s_1 = 1$ , thus explaining why the errors are small for the experiments for which the grid slope ratio is close to one (e.g., for the 64x24 and 128x48 grids in the LARGE case). As far as the biharmonic operator is concerned, the results are very consistent with the one obtained with the rotated laplacian (Fig. 14). The SW-TRIADS performs better than the TRIADS when the grid slope ratio exceeds one, which is an important asset when considering grid cells with a small aspect ratio  $\Delta x_3/\Delta x_1$  as it can happen for a  $\sigma$ -coordinate models in shallow areas. For very large values of  $s_1$  the direction of diffusion becomes less and less accurate ( Fig. 13 j), Fig. 14 h)) and a larger vertical stencil should potentially be required to lessen the discretization errors.

Monotonicity violations are obvious from Fig. 13 and 14 where the blue-shaded areas correspond to

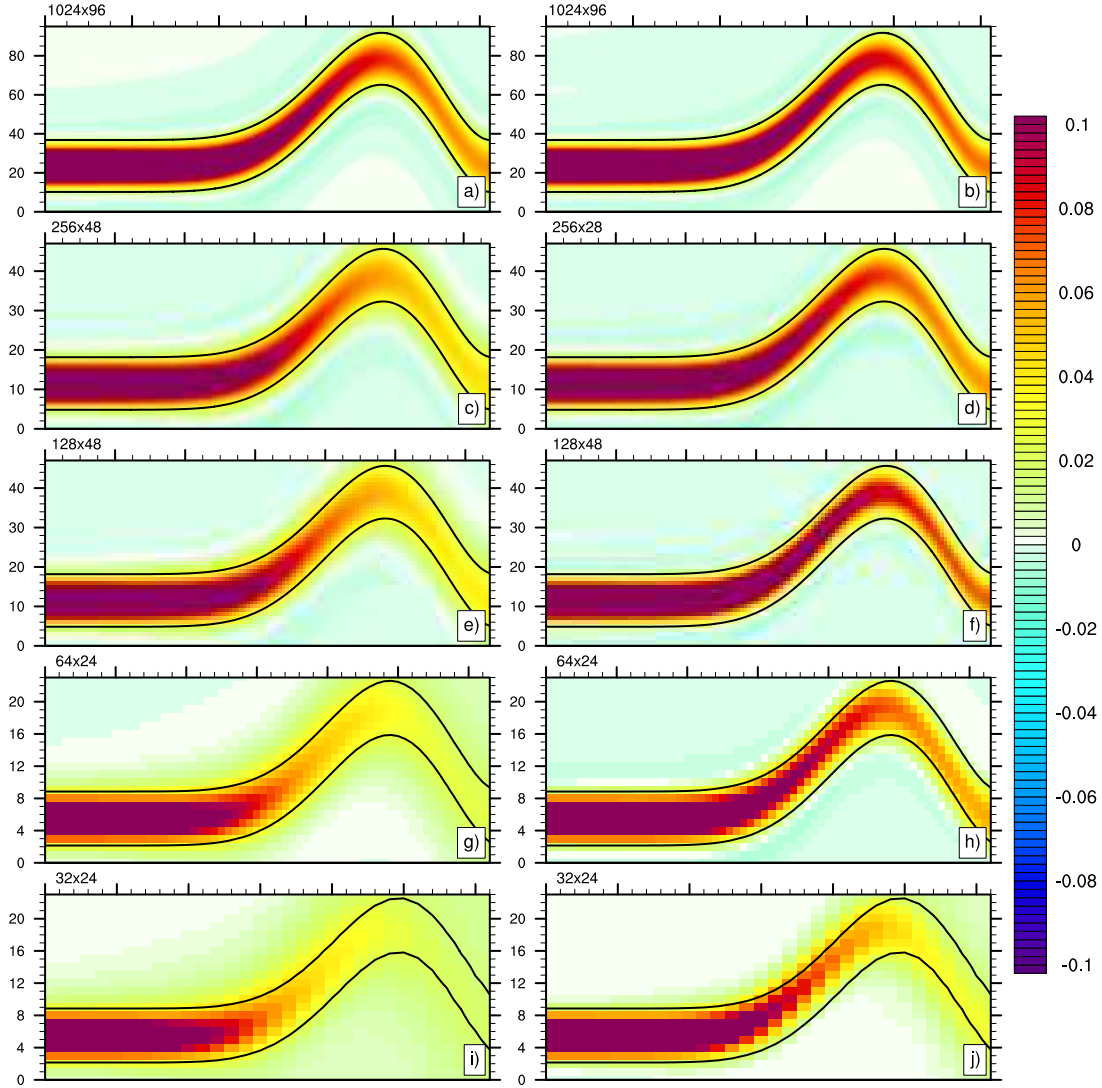


Fig. 13. Snapshots of a passive tracer diffused along sloping isopycnals after  $25 \times 10^{-3}$  s of integration in the LARGE case. Results obtained with the TRIADS scheme (resp. SW-TRIADS scheme) are shown in the left panels a), c), e), g) and i) (resp. the right panels b), d), f), h) and j)) for different grid resolutions.

negative values whereas the initial condition has positive values only. When the discretization errors are increasing, thus leading to a spurious diapycnal diffusion, the false extrema are significantly moderated. The biharmonic generates large negative values (Fig. 14), however those min-max violations are not necessarily associated with the rotation of the tensor because the horizontal biharmonic also produces overshoots with a very similar order of magnitude (not shown).

### 6.3 Overshootings

Mathieu and Deleersnijder (1998) emphasized the non-monotonic behavior of the discretized rotated diffusion operator. This can be seen simply by looking at the diffusion of a dirac signal along a direction non-aligned with the computational grid (Fig. 15). Negative bands are generated on both sides of the isopycnal along with the Dirac signal is propagating. The rotated biharmonic presents two bands of neg-

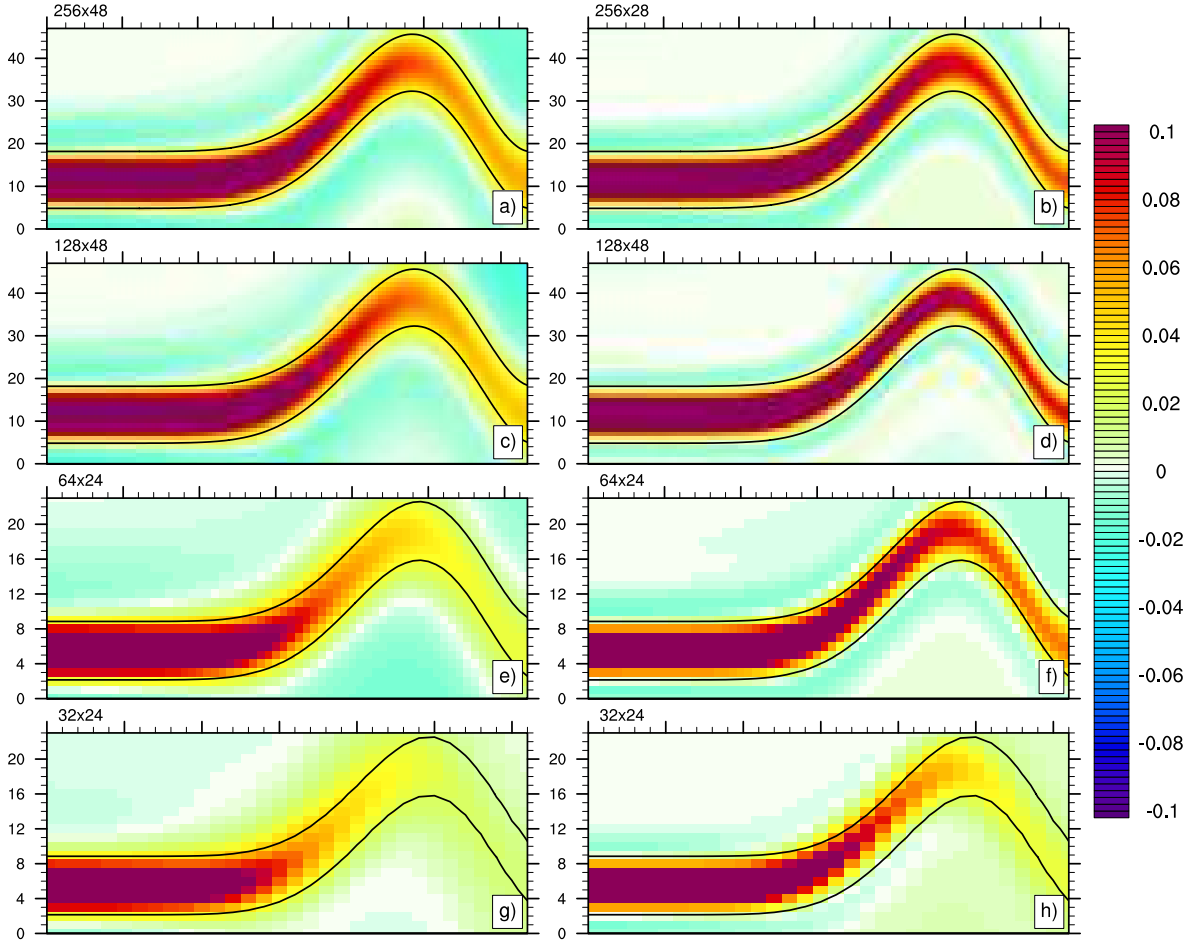


Fig. 14. Snapshots of a passive tracer diffused with a biharmonic operator along sloping isopycnals after  $25 \times 10^{-3}$  s of integration in the LARGE case. Results obtained with the TRIADS scheme (resp. SW-TRIADS scheme) are shown in the left panels a), c), e), and g) (resp. the right panels b), d), f) and h)) for different grid resolutions.

ative values separated by positive values. We can also see from those figures that the cancellation of the contribution of the cross-isopycnal points to the discretization stencil in the SW-TRIADS scheme leads to a thin signal elongated in the neutral direction while the TRIADS scheme produces a thicker signal in both the  $s_1 = 1/2$  and  $s_1 = 2$  cases. Note that the case  $s_1 = 1/2$  corresponds to a maximum of discretization errors for the SW-TRIADS scheme in the limit that  $s_1 \leq 1$ ; see discussion in Sec. 6.4.

Monotonicity is verified at a position  $(i, k)$  only if all the coefficients in the discretization stencil of the operator under consideration are positive for points others than  $(i, k)$ . It can be seen from (2.19) that neither the TRIADS, nor the SW-TRIADS scheme verifies this property. The continuous form of the rotated laplacian operator has however the property not to amplify existing extrema (Mathieu and Deleersnijder, 1998), this property is lost at the discrete level due to the discretization of the cross-derivatives. We quantify the monotonicity violation associated with the SW-TRIADS and TRIADS schemes by computing

$$\varepsilon_{i,k} = \max(q_{i,k}^{n+1} - q_{i,k}^{\max}, 0) - \min(q_{i,k}^{n+1} - q_{i,k}^{\min}, 0) \quad (6.4)$$

where

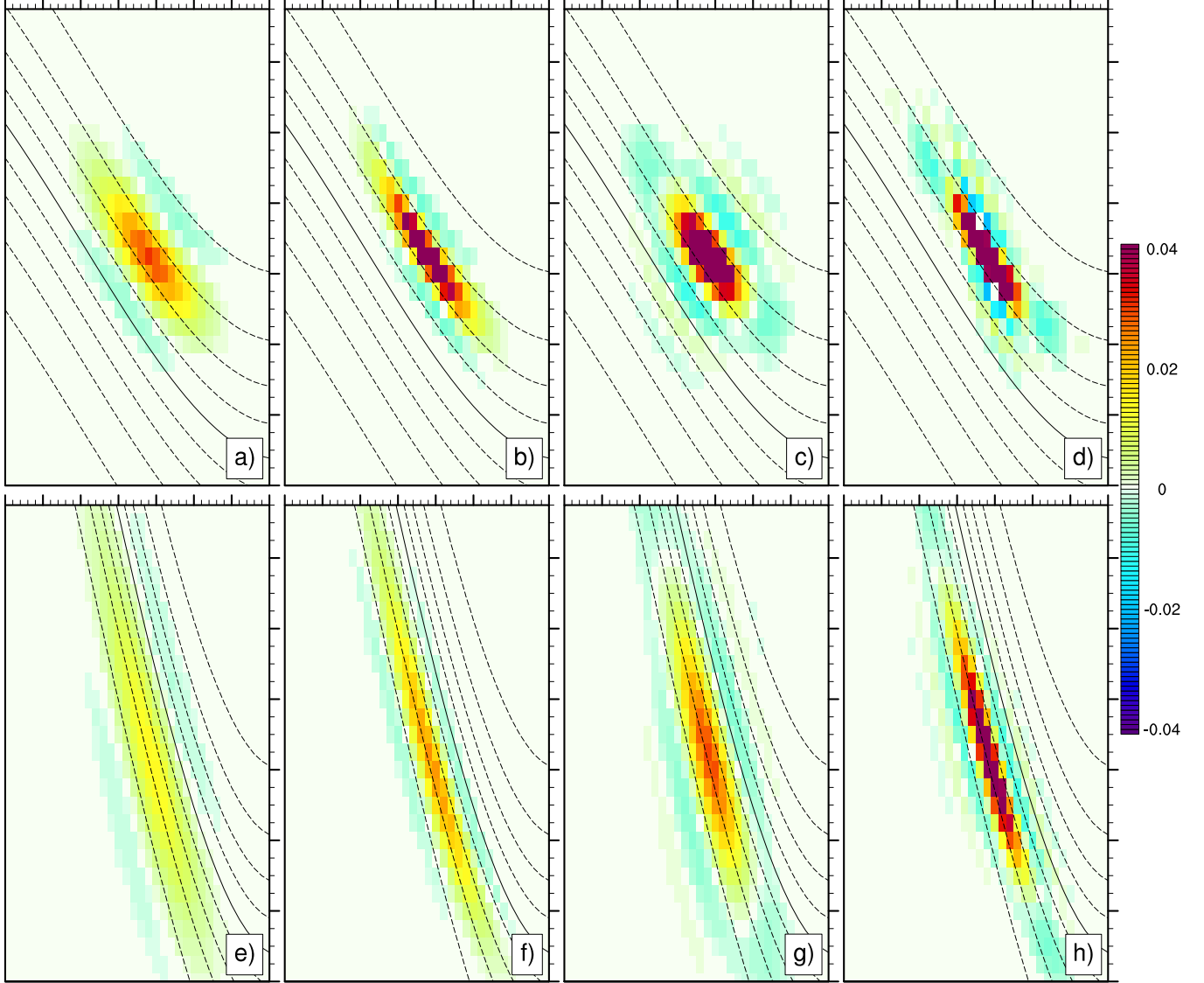


Fig. 15. Diffusion of a Dirac signal along sloping isopycnals (contours) after  $t = 2.5 \times 10^{-4}$  s : a) and e) with a rotated laplacian using the TRIADS discretization; b) and f) with a rotated laplacian using the SW-TRIADS discretization; c) and g) with a rotated biharmonic using the TRIADS discretization; d) and h) with a rotated biharmonic using the SW-TRIADS discretization. The tracer field is deliberately not smoothed, cell-averaged values are shown. Note that blue-shaded areas indicate negative values. The grid slope ratio is  $s_1 \approx 1/2$  for panels a),b),c) and d), and  $s_1 \approx 2$  for e),f),g) and h).

$$q_{i,k}^{(\min)} = \min(q_{i\pm c,k\pm c}^n), \quad q_{i,k}^{(\max)} = \max(q_{i\pm c,k\pm c}^n), \quad (6.5)$$

with  $c = 1$  for the laplacian operator and  $c = 2$  for the biharmonic. The evolution of the maximum value of  $\varepsilon_{i,k}$  is shown in Fig. 16 for the diffusion of a Dirac signal. This experiment is done on the  $256 \times 48$  grid a diffusivity  $\kappa_1$  and an hyperdiffusivity  $B_1$  chosen so that it provides the same dissipative time scale at the scale of the grid (i.e.,  $B_1 = \kappa_1 \Delta x^2$ ).

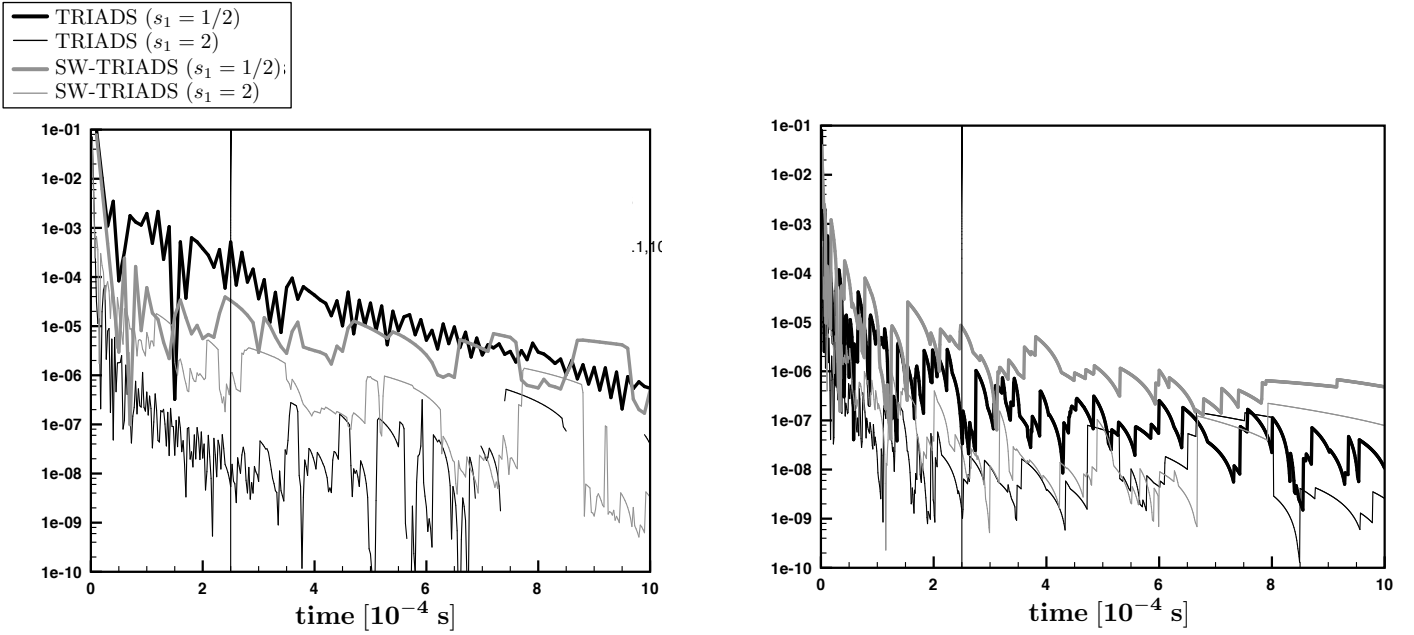


Fig. 16. Magnitude of min-max violations  $\varepsilon$  for the rotated laplacian (left) and for the rotated biharmonic (right) discretized either with the TRIADS scheme (blue and red lines) or the SW-TRIADS scheme (purple and orange lines). The vertical line at  $t = 2.5 \times 10^{-4}$  s represents the time corresponding to the plots in Fig. 15.

Fig. 16 shows that the magnitude of the min-max violations is of the same order for the rotated harmonic and biharmonic operators. Even if Fig. 15 suggests that larger negative values are produced by the biharmonic operator, this does not mean that the min-max violations are larger in this case, as we see from Fig. 16. This is explained by the fact that the amount of diapycnal diffusion, associated with discretization errors, is smaller with the rotated biharmonic than with the rotated laplacian and is insufficient to moderate the false extrema.

#### 6.4 SW-TRIADS-COMBI Scheme and Scaling of the Discretization Errors

In this paragraph, we borrow the concept of the so-called COMBI scheme, introduced in Beckers et al. (2000), to assess the evolution of the discretization errors of the SW-TRIADS scheme with respect to  $s_m$ . To minimize overshootings, Beckers et al. (2000) propose an inconsistent COMBI scheme which consists in adding an additional horizontal or vertical diffusion to cancel the negative weights in the stencil (2.19) and thus to maintain monotonicity. With the TRIADS discretization this is not straightforward because the negative weights lies systematically in the diagonal direction. For the SW-TRIADS scheme, it is easier to apply the concept of the COMBI scheme; in (2.19) we see that, for  $s_m$  positive and  $s_m < 1$ , it is sufficient to add a vertical diffusion with a coefficient proportional to  $s_m(1 - s_m)$  and, for  $s_m > 1$ , it is sufficient to add an horizontal diffusion with a coefficient proportional to  $s_m - 1$ . In Fig. 17, we show the diffusion of a Dirac signal using the SW-TRIADS-COMBI scheme (defined as the SW-TRIADS scheme combined with the COMBI approach). The scheme seems to behave relatively well for  $s_1 \approx 2$ , however the amount of spurious diapycnal diffusion is larger compared with the SW-TRIADS scheme. For  $s_1 \approx 1/2$ , the orientation of the diffusion is less accurate and the dirac signal is propagated too vertically relative to the neutral direction. We can show that the extra diffusivity  $\kappa_v$  used for  $s_m < 1$  in the SW-TRIADS-COMBI scheme is



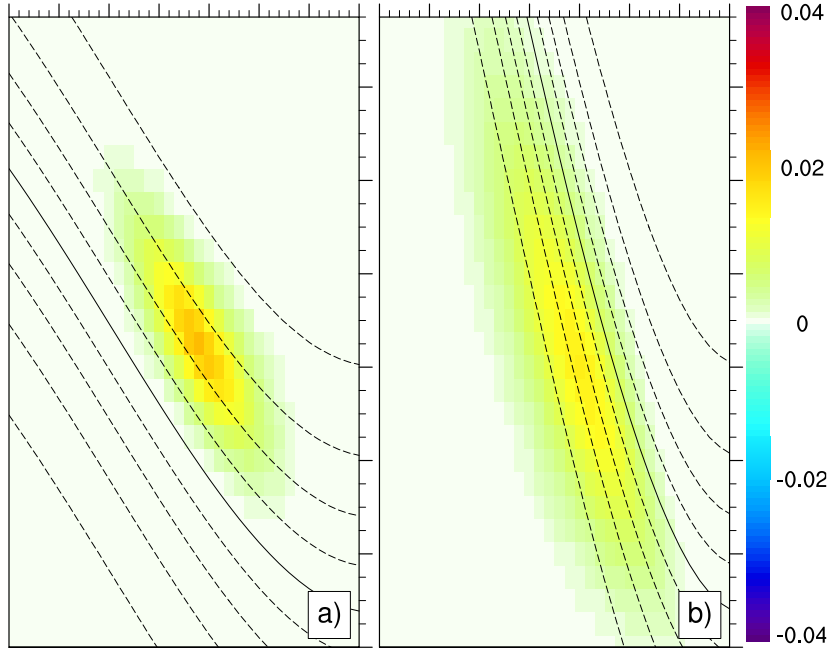


Fig. 17. Diffusion of a Dirac signal (shaded) along sloping isopycnals (contours) using the SW-TRIADS-COMBI scheme : a) for  $s_1 \approx 1/2$ ; b) for  $s_1 \approx 2$ . The results are shown after  $2.5 \times 10^{-4}$  s of integration. Panel a) should be compared with Fig. 15, b) and panel b) with with Fig. 15, f).

$$\kappa_v = \kappa_m \left( \frac{\Delta x_3}{\Delta x_m} |\alpha_m| - \alpha_m^2 \right).$$

The expression of  $\kappa_v$  is also useful to quantify the discretization errors associated with the SW-TRIADS scheme. The minimum of the error is obtained for  $\alpha_m = 0$ , or  $\alpha_m = \Delta x_3 / \Delta x_m$  (i.e.,  $s_m = 1$ ) because  $\kappa_v$  cancels in those cases, hence the scheme reduces to the classical 1,-2,1 stencil and is monotonic. The maximum of the discretization error is for  $\alpha_m = \Delta x_3 / 2 \Delta x_m$  (i.e.,  $s_m = 1/2$ ), which leads to  $\kappa_v^{\max} = \frac{\kappa_m}{4} \left( \frac{\Delta x_3}{\Delta x_m} \right)^2$ . For  $s_m > 1$ , the discretization error scales with  $s_m - 1$  (which corresponds to the background horizontal diffusivity of the SW-TRIADS-COMBI scheme) and thus keeps increasing with increasing values of  $s_m$ .

The value of  $\kappa_v^{\max}$  can be compared with the background vertical diffusivity associated with the vertical mixing parameterizations currently in use in ocean models, which is of the order of  $10^{-5} \text{m}^2 \text{s}^{-1}$  below the mixed layer. We see that it is not unusual, regarding the typical horizontal and vertical resolutions used by climate models, that the background vertical diffusion in the oceanic interior is sufficient to significantly moderate the overshootings associated with the use of a rotated mixing operator discretized with the SW-TRIADS scheme, as long as the grid slope ratio is smaller than one.

## Conclusion

The use of a mixing operator non-aligned with the computational grid raises difficulties to maintain several properties of the non-rotated operators : numerical stability, strict global tracer variance dissipation, monotonicity (in the case of the laplacian), and accuracy when the slopes steepen to greater than the grid aspect ratio. In this paper, we present a set of conservative space-time discretizations to investigate those delicacies. Since we do not want the numerical models to be penalized by diffusive processes, we consider only linear spatial discretizations and first-order accurate temporal integrations. We do not strive to design higher order schemes because the diffusive terms are used either for numerical or physical closure (or both) and are thus supposed to vanish anyway for sufficiently high resolution.

In the same spirit as the methods commonly used for multi-dimensional parabolic problems with mixed derivatives, we introduce a simple time-stepping scheme and we show its relevance to handle the time-integration of the rotated harmonic and biharmonic operators. This method, which is known as Method of Stabilizing Corrections (MSC), consists of a multi-stage approach : at the first stage, a consistent (explicit) approximation of the rotated operator is evaluated, while all succeeding stages serve only to improve the stability. For both the rotated harmonic and biharmonic operators, the stabilizing step is done using a vertical laplacian whose diffusivity is chosen through linear stability analysis. The proposed scheme is made to enable the rotated operators to be advanced with the same time step than the non-rotated ones. For large grid slope ratios, the scheme can however suffer from a lack of damping of the smallest resolved scales in the vertical. To alleviate this problem a time-splitting of the diffusion could be used with a few (typically of the order of 4) small time steps using the (MSC) scheme within a larger baroclinic time step. This approach would still be significantly more computationally efficient than using an explicit Euler scheme.

As noted previously in Beckers et al. (2000), all the consistent linear spatial discretizations of the rotated operators produce false diapycnal mixing and min-max violations. This issue is cumbersome because, in general, min-max violations are larger for schemes with small inherent diapycnal mixing, and conversely schemes with a larger diapycnal mixing associated with discretization errors tend to moderate the overshootings. The rotated biharmonic does not seem to generate significantly larger monotonicity violations than the rotated laplacian and than the non-rotated biharmonic. Furthermore, we show that different spatial discretizations of the rotated operators can lead to a quite different behavior of the solution especially for large grid slope ratios. A compact discretization (referred as to SW-TRIADS) whose stencil adapts to the orientation of the slope provides more accuracy, hence has less spurious diapycnal mixing, for situations where the isoneutral slope is greater than the vertical to horizontal aspect ratio. However, this discretization does not damp very efficiently small-scale noise, and dispersive errors could thus go uncontrolled with this approach. As shown in Beckers et al. (2000), the SW-TRIADS scheme can easily be made monotonic by adding a vertical background diffusion (as long as the grid slope ratio is less than one) at the expense of a larger dianeutral contribution. However, under some circumstances, the order of magnitude of this background diffusivity is expected to be of the same order as the physically admissible dianeutral diffusivity.

We show that rotated operators must be used with care if we rely on those operators to efficiently smooth grid-scale noise along the computational grid. Adding a background vertical, as it is commonly done through the vertical mixing parameterization, or horizontal diffusion is a way to erase the deficiencies in terms of smoothing but at the expense of additional numerically-induced spurious diapycnal mixing. The requirements in terms of noise control are tied to the numerical schemes used to handle the

tracer advection. When low-order schemes are used, larger dispersive errors are expected and an operator providing an efficient numerical filtering is crucial. In this case, the TRIADS scheme would be recommended. An unexplored possibility could be to formulate the advection in an isopycnal framework hence the numerical noise would be prominently in the neutral direction and rotated operators would be more adequate to control this noise.

Besides the idealized configurations studied in the present paper, it is shown in Lemarié et al. (2012) that a rotated biharmonic advanced with the (MSC) scheme in time and using the SW-TRIADS discretization performs well for a realistic coarse-resolution basin-scale simulation with the Regional Oceanic Modelling System (ROMS, Shchepetkin and McWilliams, 2005)  $\sigma$ -coordinate model. Using this space-time discretization, we have demonstrated that the use of an isoneutral biharmonic operator significantly improves the conservation of intermediate water properties and leads to a tightening of the thermocline compared to an iso- $\sigma$  biharmonic operator. Because the grid slope ratio with a  $\sigma$ -coordinate is generally larger than with a  $z$ -level model we took advantage of the gain of accuracy provided by the SW-TRIADS scheme.

An alternative approach to the use of rotated mixing operators is the design of a vertical coordinate system following the isopycnals (Hallberg and Adcroft, 2009; Leclair and Madec, 2011). The present paper is a complementary effort in exploring the merits of different approaches to representing the nearly adiabatic flow in the oceanic interior.

## Acknowledgements

Support was provided by the Office of Naval Research, Grant N00014-08-1-0597.

## References

- Andreev, V.B., 1967. Splitting operator difference schemes for general second order p-dimensional parabolic equations with mixed derivatives. U.S.S.R. Comput. Math. Math. Phys 7, 92 – 104.
- Beckers, J.M., Burchard, H., Campin, J.M., Deleersnijder, E., Mathieu, P.P., 1998. Another reason why simple discretizations of rotated diffusion operators cause problems in ocean models: Comments on “Isonneutral diffusion in a  $z$ -coordinate ocean model”. J. Phys. Oceanogr. 28, 1552–1559.
- Beckers, J.M., Burchard, H., Deleersnijder, E., Mathieu, P.P., 2000. Numerical discretization of rotated diffusion operators in ocean models. Mon. Weather Rev. 128, 2711–2733.
- Colas, F., Capet, X., McWilliams, J.C., Li, Z., 2012. Mesoscale eddy buoyancy flux and eddy-induced circulation in eastern-boundary upwelling systems. J. Phys. Oceanogr. Under review.
- Cox, M.D., 1987. Isopycnal diffusion in a  $z$ -coordinate ocean model. Ocean Modelling, (unpublished manuscripts) 74, 1–5.
- Craig, I., Sneyd, A., 1988. An alternating-direction implicit scheme for parabolic equations with mixed derivatives. Comput. Math. Applic. 16, 341–350.
- Cushman-Roisin, B., Beckers, J.M., 2011. Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects. Academic Press.
- Danabasoglu, G., McWilliams, J.C., 1995. Sensitivity of the global ocean circulation to parameterizations of mesoscale tracer transports. J. Climate 8, 2967–2987.

- Danabasoglu, G., McWilliams, J.C., Gent, P.R., 1994. The role of mesoscale tracer transports in the global ocean circulation. *Science* 264, 1123–1126.
- Delhez, E.J., Deleersnijder, E., 2007. Overshootings and spurious oscillations caused by biharmonic mixing. *Ocean Modell.* 17, 183 – 198.
- Douglas, J., 1962. Alternating direction methods for three space variables. *Numer. Math.* 4, 41–63.
- Douglas, J., Gunn, J.E., 1964. A general formulation of alternating direction methods. *Numer. Math.* 6, 428–453.
- Ferrari, R., McWilliams, J.C., Canuto, V., Dubovikov, M., 2008. Parameterization of eddy fluxes near oceanic boundaries. *J. Climate* 21, 2770–2789.
- Fox-Kemper, B., Ferrari, R., Hallberg, R., 2008. Parameterization of mixed layer eddies. part i: Theory and diagnosis. *J. Phys. Oceanogr.* 38, 1145–1165.
- Gent, P.R., 2011. The Gent-McWilliams parameterization: 20/20 hindsight. *Ocean Modell.* 39, 2 – 9.
- Gent, P.R., McWilliams, J.C., 1990. Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.* 20, 150–155.
- Gnanadesikan, A., Griffies, S.M., Samuels, B.L., 2007. Effects in a climate model of slope tapering in neutral physics schemes. *Ocean Modell.* 16, 1–16.
- Goosse, H., Campin, J.M., Deleersnijder, E., Fichet, T., Mathieu, P.P., Morales Maqueda, M.A., Tartiville, B., 2008. Description of the CLIO model version 3.0, 7 December 2008 <http://www.astr.ucl.ac.be/research/CLIO/clio30.pdf>.
- Griffies, S.M., 2004. *Fundamentals of Ocean Climate Models*. Princeton University Press, Princeton, NJ, USA.
- Griffies, S.M., 2010. Elements of MOM4p1, in: GFDL Ocean Group Technical Report No. 6. NOAA/GFDL, February 19, 2010. 444 pp.
- Griffies, S.M., Gnanadesikan, A., Pacanowski, R.C., Larichev, V., Dukowicz, J.K., Smith, R.D., 1998. Isonutral diffusion in a z-coordinate ocean model. *J. Phys. Oceanogr.* 28, 805–830.
- Hallberg, R., Adcroft, A., 2009. Reconciling estimates of the free surface height in lagrangian vertical coordinate ocean models with mode-split time stepping. *Ocean Modell.* 29, 15 – 26.
- Hecht, M.W., 2010. Cautionary tales of persistent accumulation of numerical error: Dispersive centered advection. *Ocean Modell.* 35, 270–276.
- in 't Hout, K., Welfert, B., 2009. Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms. *Appl. Numer. Math.* 59, 677 – 692.
- van der Houwen, P., Verwer, J., 1979. One-step splitting methods for semi-discrete parabolic equations. *Computing* 22, 291–309.
- Hundsdorfer, W., 2002. Accuracy and stability of splitting with stabilizing corrections. *Appl. Numer. Math.* 42, 213 – 233.
- Leclair, M., Madec, G., 2011.  $\tilde{z}$ -coordinate, an arbitrary lagrangian-eulerian coordinate separating high and low frequency motions. *Ocean Modell.* 37, 139 – 152.
- Lemarié, F., Kurian, J., Shchepetkin, A.F., Molemaker, M.J., Colas, F., McWilliams, J.C., 2012. Are There Inescapable Issues Prohibiting the use of Terrain-Following Coordinates in Climate Models ? *Ocean Modell.* 42, 57–79.
- Lengaigne, M., Madec, G., Menkes, C., 2003. Impact of isopycnal mixing on the tropical ocean circulation. *J. Geophys. Res.* 108.
- Madec, G., 2008. NEMO ocean engine, in: Note du Pole de modélisation No. 27. Institut Pierre-Simon Laplace (IPSL), France.
- Manfredi, G., Ottaviani, M., 1999. Finite-difference schemes for the diffusion equation, in: Leach, P.,

- Bouquet, S., Rouet, J.L., Fijalkow, E. (Eds.), *Dynamical Systems, Plasmas and Gravitation*. Springer Berlin / Heidelberg. volume 518 of *Lecture Notes in Physics*, pp. 82–92.
- Marchesiello, P., Debreu, L., Couvelard, X., 2009. Spurious diapycnal mixing in terrain-following coordinate models: The problem and a solution. *Ocean Modell.* 26, 159–169.
- Mathieu, P.P., Deleersnijder, E., 1998. What is wrong with isopycnal diffusion in world ocean models? *Appl. Math. Modelling* 22, 367 – 378.
- Mathieu, P.P., Deleersnijder, E., Beckers, J.M., 1999. Accuracy and stability of the discretised isopycnal-mixing equation. *Appl. Math. Lett.* 12, 81–88.
- Redi, M.H., 1982. Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.* 12, 1154–1158.
- Shchepetkin, A., McWilliams, J.C., 2005. The Regional Oceanic Modeling System: A split-explicit, free-surface, topography-following-coordinate ocean model. *Ocean Modell.* 9, 347–404.
- Smith, R.D., Gent, P.R., 2004. Anisotropic Gent-McWilliams Parameterization for Ocean Models. *J. Phys. Oceanogr.* 34, 2541–2564.
- Smith, R.D., Jones, P., Briegleb, B., Bryan, F., Danabasoglu, G., Dennis, J., Dukowicz, J., Eden, C., Fox-Kemper, B., Gent, P., Hecht, M., Jayne, S., Jochum, M., Large, W., Lindsay, K., Maltrud, M., Norton, N., Peacock, S., Vertenstein, M., Yeager, S., 2010. The Parallel Ocean Program (POP) Reference Manual, 23 March 2010, LAUR-10-01853.
- Solomon, H., 1971. On the representation of isentropic mixing in ocean circulation models. *J. Phys. Oceanogr.* 1, 233–234.

## A Upper Bound on $\varphi_m^\theta = \varphi_m - 2\theta y_{m3}^2$

The aim of this appendix is to derive an upper bound for the function  $\varphi_m^\theta$  introduced in Sec. 3. We keep the notations previously used in the paper, and for convenience we introduce the parameters  $X_m = \tan(\phi_m/2)$ ,  $X_m \in \mathbb{R}$  ( $m = 1, 3$ ); so that the expression of  $\varphi_m^\theta$  transforms to

$$\varphi_m^\theta = 2\sigma_m \left( \frac{[(X_m + s_m X_3)^2 + (1 + s_m^2)X_m^2 X_3^2]}{(1 + X_m^2)(1 + X_3^2)} - 2\theta s_m^2 \frac{X_3^2}{1 + X_3^2} \right). \quad (\text{A.1})$$

In the remaining of this section we successively study the behavior of  $\varphi_m^\theta$  with respect to  $X_m$  and then to  $X_3$ . We first get

$$\text{Sign} \left( \frac{\partial \varphi_m^\theta}{\partial X_m} \right) = \text{Sign} \left( s_m X_3 (1 - X_m^2) + X_m (1 + X_3^2) \right), \quad (\text{A.2})$$

where the roots of the second order polynomial in  $X_m$  are

$$X_m^{(\pm)} = \frac{1 + X_3^2 \pm \sqrt{(1 + X_3^2)^2 + (2s_m X_3)^2}}{2s_m X_3}. \quad (\text{A.3})$$

$\frac{\partial \varphi_m^\theta}{\partial X_m}$  is positive for  $X_m^{(-)} \leq X_m \leq X_m^{(+)}$  which is sufficient to show that  $\varphi_m^\theta$  reaches its maximum either for  $X_m = X_m^{(+)}$  or  $X_m \rightarrow -\infty$ . In (A.1), we see that the asymptotic limit of  $\varphi_m^\theta$  for  $X_m$  going to infinity either by positive or negative values is the same; the maximum of  $\varphi_m^\theta$  is therefore obtained for  $X_m = X_m^{(+)}$ . Substituting  $X_m$  by  $X_m^{(+)}$  in (A.1) leads to

$$\varphi_m^\theta \leq \varphi_m^\theta(X_m^{(+)}) = \bar{\varphi}_m^\theta = 2\sigma_m \left( 1 + 2s_m^2(1 - 2\theta) \frac{X_3^2}{1 + X_3^2} + \sqrt{1 + 4s_m^2 \frac{X_3^2}{1 + X_3^2} \left( 1 - \frac{X_3^2}{1 + X_3^2} \right)} \right). \quad (\text{A.4})$$

After some algebra, we find

$$\text{Sign} \left( \frac{\partial \bar{\varphi}_m^\theta}{\partial X_3} \right) = \text{Sign} \left( (1 - 2\theta) \sqrt{1 + 4s_m^2 \tilde{X}_3 (1 - \tilde{X}_3)} + (1 - 2\tilde{X}_3) \right), \quad \tilde{X}_3 = \frac{X_3^2}{1 + X_3^2} \in [0, 1]. \quad (\text{A.5})$$

This derivative cancels for

$$\tilde{X}_3 = \tilde{X}_3^{(\pm)} = \frac{1}{2} \left[ 1 \pm \sqrt{\frac{(1 - 2\theta)^2 (1 + s_m^2)}{1 + (1 - 2\theta)^2 s_m^2}} \right]. \quad (\text{A.6})$$

The maximum of  $\bar{\varphi}_m^\theta$  is reached for  $\tilde{X}_3 = \tilde{X}_3^{(+)}$  if  $\theta \leq 1/2$  and  $\tilde{X}_3 = \tilde{X}_3^{(-)}$  otherwise. In both cases the maximum is the same and provides the following upper bound for  $\varphi_m^\theta$

$$\varphi_m^\theta \leq 2\sigma_m \cdot \mathcal{M}(\theta, s_m), \quad (\text{A.7})$$

where

$$\mathcal{M}(\theta, s_m) = 1 + s_m^2(1 - 2\theta) + \sqrt{(1 + s_m^2)(1 + s_m^2(1 - 2\theta)^2)}. \quad (\text{A.8})$$

## B Stability Analysis With the SW-TRIADS Scheme

In this section, we extend the results found in Sec. 3 and 4 for the TRIADS scheme to the SW-TRIADS discretization. The function  $\varphi_m^\theta$  (with  $m = 1, 2$ ) defined in (3.21) for  $\beta_m = 0$  (i.e., the TRIADS scheme) becomes

$$\varphi_m^{\theta, \text{SW}} = 2\sigma_m \left( \frac{[(X_m + s_m X_3)^2 + (s_m - 1)^2 X_m^2 X_3^2]}{(1 + X_m^2)(1 + X_3^2)} - 2\theta s_m^2 \frac{X_3^2}{1 + X_3^2} \right) \quad (\text{B.1})$$

for  $\beta_m = 1$  (i.e., the SW-TRIADS scheme with  $s_m \geq 0$ ). In this appendix, we consider only the case  $s_m \geq 0$ , it is straightforward to extend by symmetry the results to the case  $s_m \leq 0$ . As we did in App. A for  $\varphi_m^\theta$ , we study here the behavior of  $\varphi_m^{\theta, \text{SW}}$  with respect to  $X_m$  and  $X_3$ . After some algebra, we get

$$\text{Sign} \left( \frac{\partial \varphi_m^{\theta, \text{SW}}}{\partial X_m} \right) = \text{Sign} \left( s_m X_3 (1 - X_m^2) + X_m (1 + X_3^2 (1 - 2s_m)) \right), \quad (\text{B.2})$$

where the roots of the second order polynomial in  $X_m$  are

$$X_m^{(\pm)} = \frac{1 + X_3^2 (1 - 2s_m) \pm \sqrt{(1 + X_3^2 (1 - 2s_m))^2 + (2s_m X_3)^2}}{2s_m X_3}. \quad (\text{B.3})$$

We can easily show that the maximum of  $\varphi_m^{\theta, \text{SW}}$  is for  $X_m = X_m^{(+)}$  which leads to

$$\varphi_m^{\theta, \text{SW}} \leq \varphi_m^{\theta, \text{SW}}(X_m^{(+)}) = \overline{\varphi}_m^{\theta, \text{SW}} = 2\sigma_m \left( 1 + 2(s_m(s_m - 1) - 2\theta s_m^2) \frac{X_3^2}{1 + X_3^2} + \sqrt{1 + 4s_m(s_m - 1) \frac{X_3^2}{1 + X_3^2}} \right). \quad (\text{B.4})$$

The sign of the derivative of this function with respect to  $X_3$  reads

$$\text{Sign} \left( \frac{\partial \overline{\varphi}_m^{\theta, \text{SW}}}{\partial X_3} \right) = \text{Sign} \left( (s_m - 1) \left( 1 + \sqrt{1 + 4s_m(s_m - 1) \widetilde{X}_3} \right) - 2s_m \theta \sqrt{1 + 4s_m(s_m - 1) \widetilde{X}_3} \right) \quad (\text{B.5})$$

where  $\widetilde{X}_3 = X_3^2 / (1 + X_3^2)$ .

### Rotated Laplacian

We recall that the space-time discretization under consideration is stable for  $\varphi_1^{\theta, \text{SW}} + \varphi_2^{\theta, \text{SW}} \leq 2$ . In (B.5), we see that for  $s_m \leq 1$ ,  $\partial \overline{\varphi}_m^{\theta, \text{SW}} / \partial X_3$  is negative, hence the maximum is for  $\widetilde{X}_3 = 0$  which translates into  $\varphi_m^{\theta, \text{SW}} \leq 4\sigma_m$  in (B.4). This shows that, in this case, the scheme is stable for  $\sigma_1 + \sigma_2 \leq 1/2$  which corresponds to the stability condition of the non-rotated operator. Moreover, for  $s_m > 1$ , taking  $\theta = (s_m - 1)/s_m$  still ensures that  $\partial \overline{\varphi}_m^{\theta, \text{SW}} / \partial X_3$  is negative, and thus leads to the same stability constraint. This last remark is sufficient to conclude that the stability condition of the rotated laplacian discretized using the SW-TRIADS scheme in space and the (MSC) scheme in time is equivalent to the stability condition of the non-rotated laplacian for

$$\theta = \theta^{\text{SW}} = \max \left\{ \frac{|s_1| - 1}{|s_1|}, \frac{|s_2| - 1}{|s_2|}, 0 \right\}. \quad (\text{B.6})$$

If  $|s_1|$  and  $|s_2|$  are both smaller than 1 we get  $\theta = 0$ , and for  $|s_m| \rightarrow \infty$ ,  $\theta \rightarrow 1$  which shows that the (IMP) scheme is stable as long as the horizontal terms of the tensor are stable.

### *Rotated Biharmonic*

Using (B.6), and following the methodology used in Sec. 4.3 to derive an expression for the stabilizing diffusivity  $\tilde{\sigma}^{\text{SW}}$ , we find that

$$\tilde{\sigma} = \tilde{\sigma}^{\text{SW}} = 8 (S_1 \sigma_1 + S_2 \sigma_2) ((1 + S_1) \sigma_1 + (1 + S_2) \sigma_2), \quad S_m = \max\{s_m^2 - |s_m|, 0\}. \quad (\text{B.7})$$

Note that  $\tilde{\sigma}^{\text{SW}}$  vanishes when  $|s_1|$  and  $|s_2|$  are smaller than one.