The adaptive EVP method for solving the sea ice momentum equation

Madlen Kimmritz[a,*], Sergey Danilov[a,b], Martin Losch[a]

[a] *Alfred Wegener Institute, Bussestrasse 24, D-27570 Bremerhaven*
[b] *A. M. Obukhov Institute of Atmospheric Physics RAS, Moscow, Russia*

Abstract

Stability and convergence of the modified EVP implementation of the visco-plastic sea ice rheology by Bouillon et al., Ocean Modell., 2013, is analyzed on B- and C-grids. It is shown that the implementation on a B-grid is less restrictive with respect to stability requirements than on a C-grid. On C-grids convergence is sensitive to the discretization of the viscosities. We suggest to adaptively vary the parameters of pseudotime subcycling of the modified EVP scheme in time and space to satisfy local stability constraints. This new approach generally improves the convergence of the modified EVP scheme and hence its numerical efficiency. The performance of the new "adaptive EVP" approach is illustrated in a series of experiments with the sea ice component of the MIT general circulation model (MITgcm) that is formulated on a C-grid.

Keywords: VP rheology, EVP rheology, Sea ice, MITgcm, B-grid, C-grid, adaptive relaxation parameter

2010 MSC: 00-01, 99-00

1. Introduction

The viscous-plastic (VP) rheology (Hibler III, 1979), connecting sea ice deformation rates with ice stresses, forms the basis of most climate sea-ice models. The resulting set of equations of ice dynamics is very stiff and thus calls for the

*Corresponding author (now affiliated to Nansen Environmental and Remote Sensing Center and Bjerknes Centre for Climate Research, Bergen, Norway)

Email addresses: `madlen.kimmritz@awi.de` (Madlen Kimmritz), `sergey.danilov@awi.de` (Sergey Danilov), `martin.losch@awi.de` (Martin Losch)

design of efficient solution methods to avoid the restriction to very small time steps in standard explicit methods. Partial linearization allows the stiff part of the problem to be treated implicitly, but requires iterative solvers (Zhang and Hibler, 1997). Although this linearization lifts the time step restriction, it requires many (Picard) iterations to recover the full nonlinear solution. Traditionally only a few Picard iterations are made and convergence is sacrificed (Lemieux and Tremblay, 2009). This motivated the development of fully nonlinear Jacobian-free Newton-Krylov (JFNK) solvers (Lemieux et al., 2010, 2012, Losch et al., 2014). They converge faster than previous methods but still remain an expensive solution.

The elastic-viscous-plastic (EVP) method is an alternative to implicit methods. It relaxes the time step limitation of the explicit VP method by introducing an additional (artificial, not physically motivated) elastic term to the stress equations. This allows a fully explicit time stepping scheme with much larger time steps than possible for the VP method (Hunke and Dukowicz, 1997, Hunke, 2001), but still requires subcycling within the external time step commonly set by the ocean model. The effects of the additional elasticity term, however, are reported to lead to noticeable differences in the deformation field, and result in solutions with smaller viscosities and weaker ice (e.g., Lemieux et al., 2012, Losch et al., 2010, Losch and Danilov, 2012, Bouillon et al., 2013).

In many cases, these effects are linked to the violation of local stability limits (analogous to the Courant number constraint for advection) associated with the explicit time stepping scheme of the subcycling process (Hunke and Dukowicz, 1997, Hunke, 2001). Their most frequent manifestation is grid-scale noise in the ice velocity derivatives and hence in ice viscosities, in particular, on meshes with fine or variable resolution (Losch and Danilov, 2012) (the numerical code may remain stable and simulate smooth fields of ice concentration and thickness). In an attempt to improve the performance of the EVP method, a modification of the time-discrete model was proposed by adding an inertial time stepping term to the momentum balance (Lemieux et al., 2012). This mEVP (modified EVP) method was reformulated by Bouillon et al. (2013) as a "pseudotime" iterative

scheme. By construction, it should lead to solutions that are identical to those of the VP method provided the scheme is stable and runs to convergence. The analysis of mEVP for a simplified one-dimensional (1D) case suggests that the stability is defined by a single parameter that depends on the resolution, the time step, the ice viscosity, and on the relaxation parameters of the pseudotime stepping (Bouillon et al., 2013, Kimmritz et al., 2015),.

Although the 1D analysis is expected to be valid at least qualitatively in two dimensions (2D), there are a few aspects that are not covered by the 1D analysis: the velocity and stress divergence vectors are not collinear in 2D; velocities are staggered in space (on a C-grid) but are collocated on a B-grid, so that on a C-grid one works with normal velocity components rather than the full velocity vector (as on the B-grid); on C-grids the components of the strain rate tensor and the stress components are not collocated. These aspects affect the convergence properties of the method. Several C-grid implementations have been suggested in literature (e.g. Bouillon et al., 2013, Lemieux et al., 2012, Losch et al., 2010).

This work extends the analysis of Kimmritz et al. (2015) by exploring the impact of space discretizations on the stability properties of the mEVP method. Motivated by this analysis we propose a new adaptive EVP implementation (aEVP). In this scheme the parameters of the pseudotime stepping are locally adjusted in each pseudotime subcycle in order to ensure stability. In simple experiments we demonstrate that this scheme leads to a significant improvement of the convergence properties.

The article is organized as follows: In Section 2 we briefly review the governing equations, the mEVP scheme as formulated in Bouillon et al. (2013) and its discretization on B- and C-grids. We continue with the stability analysis of the linearized 2D equations in Section 3, and introduce the aEVP method and explore its stability properties in Section 4. In Section 5, we illustrate our results in experiments performed with the sea ice component of an ocean general circulation model (MITgcm, see the source code at *http://mitgcm.org*). Conclusions and outlook are given in Section 6.

2. Model description

The horizontal momentum balance of sea ice is written as

$$m(\partial_t + f\mathbf{k}\times)\mathbf{u} = a\boldsymbol{\tau} - C_d a\rho_o(\mathbf{u}-\mathbf{u}_o)|\mathbf{u}-\mathbf{u}_o| + \mathbf{F} - mg\nabla H. \tag{1}$$

Here m is the ice (plus snow) mass per unit area, f is the Coriolis parameter and $\mathbf{k}$ the vertical unit vector, a the ice concentration, $\mathbf{u}$ and $\mathbf{u}_o$ the ice and ocean velocities, ρ_o is the ocean water density, $\boldsymbol{\tau}$ the wind stress, H the sea surface elevation, g the acceleration due to gravity and $F_l = \partial\sigma_{kl}/\partial x_k$ the divergence of the internal stress tensor σ_{kl} (with indices k, l denoting x_1 and x_2 directions). We follow Bouillon et al. (2013) in writing the VP constitutive law as

$$\sigma_{kl}(\mathbf{u}) = \frac{P}{2(\Delta+\Delta_{min})}\left[(\dot{\epsilon}_d - \Delta)\delta_{kl} + \frac{1}{e^2}(2\dot{\epsilon}_{kl} - \dot{\epsilon}_d\delta_{kl})\right], \tag{2}$$

with

$$\dot{\epsilon}_{kl} = \frac{1}{2}\left(\partial_k u_l + \partial_l u_k\right), \quad \Delta = \left(\dot{\epsilon}_d^2 + \frac{1}{e^2}\dot{\epsilon}_s^2\right)^{1/2}. \tag{3}$$

The stress tensor $\boldsymbol{\sigma}(\mathbf{u})$ is symmetric, i.e. $\sigma_{12}(\mathbf{u}) = \sigma_{21}(\mathbf{u})$. The term $\dot{\epsilon}_d = \dot{\epsilon}_{kk}$ describes the divergence, and $\dot{\epsilon}_s = ((\dot{\epsilon}_{11} - \dot{\epsilon}_{22})^2 + 4\dot{\epsilon}_{12}^2)^{1/2}$ is the shear. The parameter $e = 2$ is the ratio of the major axes of the elliptic yield curve. Note that the use of the replacement pressure, $(\Delta/(\Delta + \Delta_{\min}))P$ (Hibler III and Ip, 1995) in the formulation of the VP constitutive law (2) ensures that the stress state is on an elliptic yield curve even when $\Delta \lesssim \Delta_{\min}$. The ice strength P is parameterized as $P = hP^* e^{-c^*(1-a)}$, where h is the mean thickness of the grid cell, and the constants P^* and c^* are set to $P^* = 27500$ Nm^{-2} and $c^* = 20$. For future reference we introduce the bulk and shear viscosities $\zeta = 0.5\,P/(\Delta + \Delta_{min})$ and $\eta = \zeta/e^2$.

2.1. The mEVP scheme as a pseudotime iterative scheme

The difficulty in integrating (1) is the stiff character of the stress term, which requires prohibitively small time steps in an explicit time stepping scheme. The traditional approach is either implicit (Zhang and Hibler, 1997) where viscosities are estimated at the previous nonlinear iteration and several iterations are made,

or explicit, through the EVP formulation (Hunke and Dukowicz, 1997, Hunke and Lipscomb, 2008) where adding a pseudo-elastic term reduces the time step limitations. A discussion of the convergence issues can be found, for instance, in Bouillon et al. (2013), Kimmritz et al. (2015) and is not repeated here.

The suggestion by Bouillon et al. (2013) is equivalent, up to details of treating the Coriolis and the ice-ocean drag terms, to formulating the mEVP method as:

$$\boldsymbol{\sigma}^{p+1} - \boldsymbol{\sigma}^{p} = \frac{1}{\alpha}\Big(\boldsymbol{\sigma}(\mathbf{u}^{p}) - \boldsymbol{\sigma}^{p}\Big), \tag{4}$$

$$\mathbf{u}^{p+1} - \mathbf{u}^{p} = \frac{1}{\beta}\Big(\frac{\Delta t}{m}\nabla\cdot\boldsymbol{\sigma}^{p+1} + \frac{\Delta t}{m}\mathbf{R}^{p+1/2} + \mathbf{u}_n - \mathbf{u}^{p}\Big). \tag{5}$$

In (5), $\mathbf{R}$ sums all the terms in the momentum equation except for the rheology and the time derivative, Δt is the external time step of the sea ice model commonly set by the ocean model, the index n labels the time levels of the model time, and the index p is that of pseudotime (subcycling step number). The Coriolis term in $\mathbf{R}^{p+1/2}$ is treated implicitly in our B-grid implementation, but is explicit on the C-grid, and the ice-ocean stress term is linearly-implicit ($C_d\rho_o|\mathbf{u}_o - \mathbf{u}^p|(\mathbf{u}_o - \mathbf{u}^{p+1})$). The term $\boldsymbol{\sigma}(\mathbf{u}^p)$ in (4) implies that the stresses are estimated by (2) based on the velocity of iteration p, and $\boldsymbol{\sigma}^p$ is the variable of the pseudotime iteration. The relaxation parameters α and β in (4) and (5) are chosen to satisfy stability constraints, see Bouillon et al. (2013), Kimmritz et al. (2015). They replace the terms $2T/\Delta t_e$ and $(\beta^*/m)(\Delta t/\Delta t_e)$, where T is the elastic damping time scale and Δt_e the subcycling time step of standard EVP formulation; the parameter β^* was introduced in Lemieux et al. (2012). If (4) and (5) are iterated to convergence, their left hand sides can be set to zero leaving the VP solution:

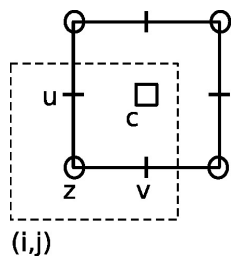
$$\frac{m}{\Delta t}\Big(\mathbf{u}_{n+1} - \mathbf{u}_n\Big) = \nabla\cdot\boldsymbol{\sigma}(\mathbf{u}_{n+1}) + \mathbf{R}^*, \tag{6}$$

with $\mathbf{R}^* = \lim_{p\to\infty}\mathbf{R}^{p+1/2}$ and $\mathbf{u}_{n+1} = \lim_{p\to\infty}\mathbf{u}^p$. While one may introduce a convergence criterion to determine the number of iteration steps, historically, the actual number of pseudotime iterations N is selected experimentally to ensure the accuracy needed. The new velocity $\mathbf{u}_{n+1}$ at time step $n+1$ is estimated at

the last pseudotime step $p = N$. The initial values for $p = 1$ are taken from the previous time step n.

2.2. Spatial discretizations

We consider discretizations on Arakawa B- and C- grids that are commonly used in sea-ice models. The positions of variables on these grids are depicted in Figure 1. Note, that in this section (i, j) is used as mesh indices. For simplicity



variables	B-grid	C-grid
scalars	c	c
velocities (u, v)	z	(u,v)
$\dot{\varepsilon}_{kk}, \sigma_{kk}$	c	c
ζ, η in bulk stress definition	c	c
$\dot{\varepsilon}_{12}, \sigma_{12}$	c	z
η in shear stress definition	c	z

Figure 1: On the left hand side the location of the cell points are sketched: c is the cell center (square symbol), z a vertex (circle), u and v the velocity points on a C-grid. All points in the dashed box are indexed with the same index pair (i, j). The table on the right hand side displays the location of the variables on B- and C-grids. Scalar quantities are ice concentration, ice mass, ice strength and sea surface elevation.

we use Cartesian coordinates and uniform grids with cell widths Δx_1 and Δx_2. The complete discretization on general orthogonal curvilinear grids can be found in Bouillon et al. (2009) and Losch et al. (2010). For convenience we introduce the notation

$$\delta_1 \phi_{i,j} = \phi_{i,j} - \phi_{i-1,j}, \qquad \delta_2 \phi_{i,j} = \phi_{i,j} - \phi_{i,j-1},$$
$$\overline{\phi_{i,j}}^{1} = (\phi_{i,j} + \phi_{i+1,j})/2, \qquad \overline{\phi_{i,j}}^{2} = (\phi_{i,j} + \phi_{i,j+1})/2$$

for a quantity ϕ at a cell with index (i, j). An expression of the form $\overline{\phi_{i,j}}^{1,2}$ defines the successive application of both directional averaging operators on ϕ. Note, that the location of the discretized derivatives depends on the respective grid arrangement of variables.

The strain rates on a B-grid are given by

$$(\dot{\varepsilon}_{11})_{ij} = \overline{\delta_1(u_1)_{i+1,j}}^{2}\Delta x_1^{-1}, \quad (\dot{\varepsilon}_{22})_{ij} = \overline{\delta_2(u_2)_{i,j+1}}^{1}\Delta x_2^{-1},$$
$$(\dot{\varepsilon}_{12})_{ij} = \frac{1}{2}\left(\overline{\delta_2(u_1)_{i,j+1}}^{1}\Delta x_2^{-1} + \overline{\delta_1(u_2)_{i+1,j}}^{2}\Delta x_1^{-1}\right) ,$$

u_1 and u_2 denote the first and the second velocity component, respectively. On C-grid, the definition of the strain rates is the same as on the B-grid but without the averaging step. In the B-grid arrangement, the divergence of the stress tensor, which contributes as a forcing in the momentum balance, is reconstructed on nodes as ($k = 1, 2$ for the two sea ice momentum equations)

$$((\nabla \cdot \boldsymbol{\sigma})_k)_{i,j} = \overline{\delta_1(\sigma_{1k})_{i,j-1}}^{2}\Delta x_1^{-1} + \overline{\delta_2(\sigma_{k2})_{i-1,j}}^{1}\Delta x_2^{-1} .$$

On a C-grid, the vector quality of the divergence is lost. Instead it is given on u and v points by

$$((\nabla \cdot \boldsymbol{\sigma})_1)_{i,j} = \delta_1(\sigma_{11})_{i,j}\Delta x_1^{-1} \quad + \delta_2(\sigma_{12})_{i,j+1}\Delta x_2^{-1} ,$$
$$((\nabla \cdot \boldsymbol{\sigma})_2)_{i,j} = \delta_1(\sigma_{12})_{i+1,j}\Delta x_1^{-1} + \delta_2(\sigma_{22})_{i,j+1}\Delta x_2^{-1} .$$

In the B-grid framework all derivatives include averaging but are collocated and share the same stencil. There is no immediate averaging of velocity derivatives for C-grid discretizations. While this results in a smaller stencil, the tensor components and derivatives are defined at different locations. For this reason we still need averaging for the determination of Δ and hence for computing the viscosities η and ζ.

Further steps in the B-grid arrangement are straightforward. On C-grids, there is some freedom in computing the viscosities. More precisely, since the bulk and shear stresses are defined at different locations, we also need to define viscosities on these different locations. We consider two options. One is introduced in Bouillon et al. (2013), the other one is the current default implementation in the sea ice component of the MITgcm (Losch et al., 2010, see the source code at *http://mitgcm.org*).

The discretization of Δ on cell centers coincides in both cases; the contributing square of the shear strain rate is formulated as a weighted average of

its adjacent nodal values. Since we treat Δx_1 and Δx_2 as constants, it reduces to $\overline{(\dot{\varepsilon}_{12}^2)}^{12}$. A formulation on more general grids can be found in Bouillon et al. (2013). The definition of the nodal shear viscosity differs in the two cases: While in Bouillon et al. (2013) it is given as the average values of the adjacent cells, the MITgcm counterpart aims to keep the stencil of the single contributions as small as possible. Denoting the former approach as C1 and the latter as C2 the shear viscosities at nodal points are given as

$$\text{(C1)} \quad \overline{\eta_{i-1,j-1}}^{12} \qquad\qquad \text{(C2)} \quad \overline{P_{i-1,j-1}}^{12} / \left(2e^2(\Delta_{ij}^z + \Delta_{min})\right)$$

with nodal value

$$\Delta_{ij}^z = \left(\overline{(\dot{\varepsilon}_{11} + \dot{\varepsilon}_{22})_{i-1,j-1}^2}^{12} + e^{-2}\left(\overline{(\dot{\varepsilon}_{11} - \dot{\varepsilon}_{22})_{i-1,j-1}^2}^{12} + 4(\dot{\varepsilon}_{12})_{ij}^2\right)\right)^{1/2}.$$

In an attempt to circumvent the ambiguity in the definition of the viscosities, we also considered an approach that first reconstructs full velocities to B-grid locations, then computes stresses and their divergence on B-grid and projects the result to the C-grid locations. Its excessive averaging and lack of commutability of derivatives, accompanied by unfavorable mathematical properties and very poor stability, however, forced us to discard it.

3. Stability analysis

We begin with generalizing the linear analysis of Kimmritz et al. (2015) to two dimensions. We will see that despite added complexity and the fact that the vectors of velocity and stress divergence are not collinear, the stability still depends on parameters that are similar to that of the 1D case and that the C-grid discretization is less stable than B-grid discretization. Similar to the 1D analysis we will assume that P and $\Delta = \Delta_{min}$ are constant, and drop $\mathbf{u}_n$ and $\mathbf{R}^{p+1/2}$ (under these assumptions C1 and C2 are similar). In order to add stability to the scheme, we take the last term $\boldsymbol{\sigma}^p$ in (4) and the last term $\boldsymbol{u}^p$ in

(5) implicitly:

$$\sigma_{kl}^{p+1} = \frac{\alpha}{\alpha+1}\sigma_{kl}^{p} + \frac{\zeta}{\alpha+1}\left((1-e^{-2})\nabla\cdot\mathbf{u}^{p}\delta_{kl} + 2e^{-2}\dot{\varepsilon}_{kl}^{p}\right)\,, \tag{7}$$

$$\mathbf{u}^{p+1} = \frac{\beta}{\beta+1}\mathbf{u}^{p} + \left(\frac{1}{\beta+1}\frac{\Delta t}{m}\right)\nabla\cdot\boldsymbol{\sigma}^{p+1}\,. \tag{8}$$

For the linear analysis we focus on a single Fourier harmonic in space

$$(\boldsymbol{\sigma}^{p}(\mathbf{x}), \boldsymbol{u}^{p}(\mathbf{x}))^{T} = \mathbf{v}_{p}e^{i\mathbf{k}\mathbf{x}} \tag{9}$$

with $(\boldsymbol{\sigma}^{p}(\mathbf{x}), \boldsymbol{u}^{p}(\mathbf{x})) = (\sigma_{11}^{p}(\mathbf{x}), \sigma_{12}^{p}(\mathbf{x}), \sigma_{22}^{p}(\mathbf{x}), u_{1}^{p}(\mathbf{x}), u_{2}^{p}(\mathbf{x}))$ and vector $\mathbf{v}_p \in \mathbb{C}^5$. After inserting expression (9) in equations (7) and (8) they reduce to a system of five equations for the components of $\mathbf{v}_p$. In matrix form, they read

$$\mathbf{v}_{p+1} = \mathrm{A}\,\mathbf{v}_{p}$$

with the 5 by 5 matrix A that corresponds to the operators on the right hand side of (7) and (8) and also incorporates the dependence on the wave vector $\mathbf{k}$. The related iterative scheme converges if $\mathbf{v}_p$ decays as p tends to infinity. Introducing the amplification factor λ as $\mathbf{v}_{p+1} = \lambda\mathbf{v}_p$, we see that such a solution is only possible if λ is an eigenvalue of the matrix A (with the eigenvector $\mathbf{v}_p$).

There are five complex-valued solutions λ_i. The formal stability condition of the discrete equations is $|\lambda_i| \leq 1$ for all $i = 1...5$. But, in analogy to the 1D case, we argue that the more restrictive condition, $|\lambda_i| < 1$ and $|\varphi_i| \ll 1$, where φ_i is the phase of λ_i, has to be imposed due to the nonlinearity of the full equations (Kimmritz et al., 2015). Because of the fifth order of the characteristic equation, we will explore the behavior of its roots numerically.

Using the notation

$$d_\sigma = \frac{\alpha}{\alpha+1}\,, \qquad d_u = \frac{\zeta}{\alpha+1}\,, \qquad c_\sigma = \frac{1}{\beta+1}\frac{\Delta t}{m}\,, \qquad c_u = \frac{\beta}{\beta+1}\,.$$

the matrix A can be written as

$$
\mathrm{A} = \begin{pmatrix} d_\sigma & 0 & 0 & e_1\psi_{x_1} & e_2\psi_{x_2} \\ 0 & d_\sigma & 0 & e_4\psi_{x_2} & e_4\psi_{x_1} \\ 0 & 0 & d_\sigma & e_2\psi_{x_1} & e_1\psi_{x_2} \\ c_\sigma d_\sigma \psi_{x_1} & c_\sigma d_\sigma \psi_{x_2} & 0 & a_1 & e_3\psi_{x_1x_2} \\ 0 & c_\sigma d_\sigma \psi_{x_1} & c_\sigma d_\sigma \psi_{x_2} & e_3\psi_{x_1x_2} & a_2 \end{pmatrix}, \tag{10}
$$

where $e_1 = d_u(1+e^{-2})$, $e_2 = d_u(1-e^{-2})$, $e_3 = c_\sigma d_u$, $e_4 = d_u e^{-2}$, and $a_l = c_u + c_\sigma d_u \left((1+e^{-2})\psi_{x_l x_l} + e^{-2}\psi_{x_{l^*}x_{l^*}}\right)$ with $l \in \{1,2\}$, $l^* = 1$ for $l = 2$ and vice versa. On a B-grid, the remaining terms (stemming from derivatives) take the form

$$
\begin{aligned}
\psi_{x_l} &= 2\mathrm{i}\sin(0.5k_l\Delta x_l)\,\cos(0.5k_{l^*}\Delta x_{l^*})/\Delta x_l, \\
\psi_{x_l x_l} &= \left(\cos(k_l\Delta x_l)\cos(k_{l^*}\Delta x_{l^*}) + \cos(k_l\Delta x_l) - \cos(k_{l^*}\Delta x_{l^*}) - 1\right)/\Delta x_{l^*}^2, \\
\psi_{x_1x_2} &= -\left(\sin(k_1\Delta x_1)\sin(k_2\Delta x_2)\right)/(\Delta x_1\,\Delta x_2).
\end{aligned}
$$

Averaging, intrinsic to the derivatives on a B-grid, leads to additional cosine multipliers, so that derivatives always depend on both components of the wave number. In contrast, on a C-grid the derivatives only depend on the wave numbers related to their directions:

$$
\begin{aligned}
\psi_{x_l} &= 2\mathrm{i}\sin(0.5k_l\Delta x_l)/\Delta x_l, \\
\psi_{x_l x_l} &= 2\left(\cos(k_l\Delta x_l) - 1\right)/\Delta x_l^2, \\
\psi_{x_1x_2} &= -4\left(\sin(0.5k_1\Delta x_1)\sin(0.5k_2\Delta x_2)\right)/(\Delta x_1\,\Delta x_2)\,.
\end{aligned}
$$

Setting either k_1 or k_2 to zero reduces the system to the 1D case where B- and C-grids coincide. Since we assumed a constant value for Δ, there is no difference between the two implementations (C1 and C2) on the C-grid.

3.1. General considerations

Throughout this section we use $\Delta x = \Delta x_1 = \Delta x_2$. Since the strongest pseudotime step limitations are expected at the largest resolved wave numbers we choose

$$
(k_1, k_2) \in \left\{ \pi\,\Delta x^{-1}\,(\cos\phi, \sin\phi) \;\middle|\; \phi \in [0,1]\cdot 2\pi \right\}. \tag{11}
$$

We set $\Delta x = 10^5$ m, $\Delta t = 3600$ s, $a = 1$, $m = 1\,\mathrm{m}$, $\Delta = 2{\cdot}10^{-7}$ s^{-1}, and $\alpha = \beta$. Figure 2 plots the eigenvalues on a B-grid and on a C-grid for $\alpha = \beta \in \{140, 500\}$ and various angles ϕ between the horizontal waves (see also (11)). In the plots we additionally depicted the unit circle in order to highlight the magnitudes and phases of the eigenvalues. In agreement with Kimmritz et al. (2015), both the magnitudes of the phases φ and the magnitudes $|\lambda|$ are controlled by α and β. The larger α and β, the closer are the eigenvalues to the stable region close to 1. There is always an eigenvalue with zero phase, which corresponds to motions

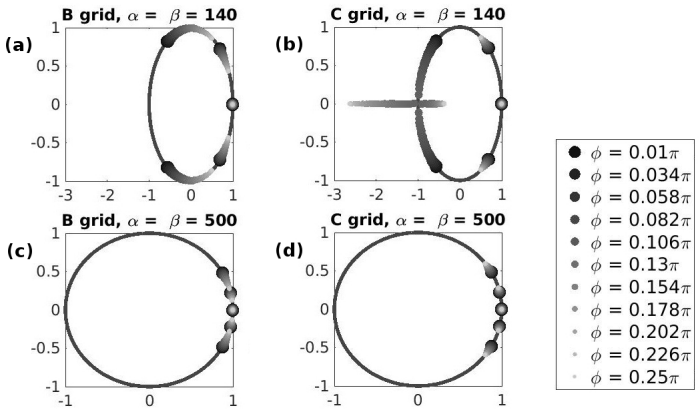


Figure 2: Eigenvalues of the system matrix $\mathbf{A}$ for the B- and the C-grid for $\alpha = \beta = 140$ (graphs (a) and (b)) and $\alpha = \beta = 500$ (graphs (c) and (d)). The wave numbers (k_1, k_2) are given by equation (11) with angle ϕ varying between 0 and $\pi/4$ with increments of 0.005. The grey circle denotes the unit circle around the origin. In the stable cases, the differences of the eigenvalues from the unit circles are $1/\alpha$, see also Table 1. For $\alpha = \beta = 140$ on C grid, the magnitudes of the eigenvalues, $|\lambda|$, exceed 1 for $\phi > 0.154\pi$.

that are little affected by the sea ice stresses. The other four eigenvalues appear in complex conjugate pairs if the solution is stable (they may become real-valued for larger Δ or smaller wave numbers). The maximum phase is larger for the eigenvalues on the C-grid indicating that the C-grid implementation is more susceptible to instability than the B-grid discretization. We assume that the additional averaging on the B-grid improves the stability of the scheme. For instance, the case $\alpha = \beta = 140$ is unstable on the C-grid, but stable on the

B-grid. At the onset of instability, two complex valued eigenvalues coincide at -1 and diverge from this point along the real axis for increasing angles ϕ. The eigenvalues in a stable situation have magnitudes of $\alpha/(1+\alpha) < 1$ (Table 1). In the numerical analysis, we observed eigenvalues with magnitudes of $\alpha/(1+\alpha)$ and $\beta/(1+\beta)$ for $\alpha \neq \beta$.

	$\alpha = \beta = 140$		$\alpha = \beta = 500$	
	$\max\{\lvert\lambda\rvert\}$	$\max\{\varphi\}$	$\max\{\lvert\lambda\rvert\}$	$\max\{\varphi\}$
B-grid	$0.993(*)$	$0.69\,\pi$	$0.998(*)$	$0.16\,\pi$
C-grid	2.638	π	$0.998(*)$	$0.20\,\pi$

Table 1: Eigenvalues for $\alpha = \beta \in \{140, 500\}$ with maximum absolute value or phase on a B-grid and on a C-grid as depicted in Figure 2. The symbol $(*)$ indicates, that all eigenvalues of the 5 times 5 matrix have the same magnitude ($\alpha/(1+\alpha)$).

Figure 3 presents the dependence of the maximum phase of the eigenvalues on the governing parameters for $\alpha = \beta = 250$. There is only a weak sensitivity of $\max\{\varphi\}$ on the ice mass m (not shown). Lower values of Δ, higher resolution in space, and higher ice concentrations lead to larger phases in the eigenvalues and thus to a less stable system in agreement with previous stability analyses (Kimmritz et al., 2015). For very fine meshes it is important to note that increasing the mesh resolution while scaling the time resolution at the same rate ($\Delta t \sim \Delta x$) makes the scheme unstable (Fig. 3(c)), but when the time step is reduced proportionally to the square of the spatial resolution ($\Delta t \sim \Delta x^2$), the scheme remains stable (Fig. 3(f)) in agreement with the stability constraint derived in Kimmritz et al. (2015). Reduced grid spacing Δx with constant time step Δt (Fig. 3(d)), which is a typical situation for models with locally refined meshes, leads to lower stability. Thus, the graphs in Figure 3 indicate a proper (i.e. stability preserving) scaling of Δt for mesh refinements or for meshes with strongly varying resolution. In all cases, the phase is slightly larger on the C-grid than on the B-grid.

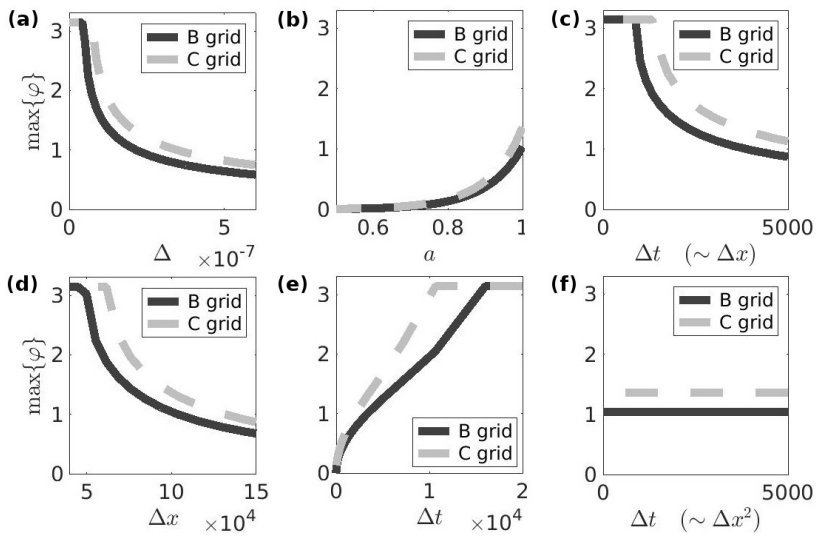


Figure 3: Dependence of the maximum phase of the eigenvalues (larger phase implies less stability) on Δ (a), on ice concentration a (b), on Δt, which scales at the same rate as Δx with initial $(\Delta t, \Delta x) = (3600\,\mathrm{s}, 10^5\,\mathrm{m})$ (c), on Δx with fixed Δt (d), on Δt with constant Δx (e) and Δt which scales with Δx^2 with initial $(\Delta t, \Delta x) = (3600\,\mathrm{s}, 10^5\,\mathrm{m})$ (f) on a B-grid (black line) and on a C-grid (grey dashed line). For small ice concentrations a, the phase is small, because the ice strength P is small.

4. The adaptive EVP method

The choice of parameters α and β is the key for providing stability of the solution. Based on the 1D analysis, Kimmritz et al. (2015) proposed to select α and β so that $\alpha\beta \gg \gamma$, where $\gamma = k^2 P \Delta t/(2\Delta m)$, with $k^2 < (\pi/\Delta x)^2$, governs stability. The regimes that are challenging for stability of the iterative process are those when γ is large and thus controls the phase (frequency) of the pseudo-time iteration. The results shown in Figure 3 and additional computations (not shown) suggest that in 2D the largest phase is controlled by the same parameter γ for a fixed wave vector direction as in the 1D case; Figure 2 also indicates that the 2D character of the problem implies some additional dependence on the wave vector direction.

Keeping α and β sufficiently large to provide stability has the downside that

the speed of convergence is slowed down and a large number of pseudotime steps N is required ($N > \alpha, \beta$) to reach convergence. In practice, very large α and β are only required in regions where viscosities ($P/2\Delta$) are large or the mesh resolution is high, while keeping them large outside of these regions only deteriorates convergence. A solution to this dilemma is making α and β variable in space and time, which is possible because mEVP, as opposed to the standard EVP approach, fully detaches α and β from the external time stepping scheme. We now introduce an approach which makes use of this possibility.

Motivated by the fact that $\gamma = k^2 P \Delta t/(2\Delta m)$ controls stability, we write it as

$$\gamma = \zeta \frac{c}{A_c} \frac{\Delta t}{m}$$

and require that $\alpha\beta \gg \gamma$. Here, A_c denotes the area of the local 2D grid cell and constant c is a numerical factor such that the term c/A_c accounts for the contribution due to the eigenvalue k^2 of the Laplacian operator, see Kimmritz et al. (2015), which has the upper limit of π^2/A_c. While this implies an upper bound of π^2 for c, c can be much smaller if ice remains smooth on the grid scale. In practice, the value of c depends on forcing, geometry of boundaries and on resolution and has to be selected experimentally. In most cases, when the solution is stable, there is no grid-scale noise so that c can be smaller than π^2 by an order of magnitude. On finer meshes the geometrical complexity of solutions may be locally increased (e.g. Losch et al., 2014), which may require using c closer to its upper bound.

In order to satisfy the stability requirement, we choose

$$\alpha = \beta = (\tilde{c}\gamma)^{1/2} \tag{12}$$

with the empirical scaling factor $\tilde{c}$. It should be sufficiently large to preserve stability, but just large enough to ensure convergence as fast as possible. The parameters c and $\tilde{c}$ can easily be combined into a single parameter, but we keep them separate here to emphasize their origin.

For instance, with $c = (0.5\pi)^2$ and $\tilde{c} = 4$, the phases of the eigenvalues, independently of the magnitudes of Δ, Δt or Δx, reach values of about 0.86π

on a C-grid and of 0.71π on a B-grid. Since the mean ice thickness enters both ice mass m and ice strength P, it has no effect on stability. Lowering the ice concentrations leads to lower maximum phases of the eigenvalues. This is due to the small exponential factor in the ice strength P for ice concentrations much smaller than 1. This factor makes γ small, so that it does not govern the behavior of the eigenvalues because the contributions from the internal stress also become small with small ice concentrations. Since α^{-1} and β^{-1} play the role of the subcycling time steps (in units of Δt), α and β should be bounded from below to ensure a sufficient accuracy of the subcycling. This adaptive approach thus guarantees stability of the iterative scheme independent of the problem parameters.

In this approach, places where α and β are large because of large values of γ will be characterized by slower convergence, but will remain stable. We suggest to select the number of pseudotime steps $N = \text{const}$ so as to provide the convergence over a dominant fraction of the domain (where γ is moderate). The convergence in local regions with high α and β will be sacrificed in favor of faster code performance. It may still be recovered over several external time steps. It is also expected that places with high α and β are those where ice velocities are small, so that incurring errors in the ice distribution are not necessarily large. If this approach is adopted, N has to be selected experimentally.

Finally, we would like to point out that the eigenvalue analysis revealed (not shown), that setting $\alpha \neq \beta$ by splitting γ in constituent multipliers generally requires an individual scaling of α and β if the resolution in time or space is varied. We do not consider this case here.

So far we were guided by the results of the linear analysis. We turn to numerical experiments to study the behavior of the adaptive EVP method in the nonlinear case.

5. Numerical experiments

In this section we explore the convergence of the full sea ice momentum equation on B- and C-grids. We will demonstrate that the discretization details of the viscosities on a C-grid influences the convergence of the mEVP method to the extent that it even may lose convergence. We will also demonstrate that the adaptive approach generally leads to improved convergence compared to simulations with constant α and β.

5.1. Experimental setup

The simple model configuration with a $L_{x_1} \times L_{x_2} = 1280\,\mathrm{km} \times 1280\,\mathrm{km}$ domain and a Cartesian grid with a constant grid size of 16 km follows that of Hunke (2001), but without topography in the model interior. The sea ice is driven by the ocean currents with the velocity (in m/s)

$$u_0 = 0.1(2x_2 - x_{2,min})/L_{x_2} \quad v_0 = -0.1(2x_1 - x_{1,min})/L_{x_1}$$

and wind stress

$$\boldsymbol{\tau} = C_a \rho_a \boldsymbol{u}_a |\boldsymbol{u}_a|$$

with atmospheric drag coefficient $C_a = 2.25 \cdot 10^{-3}$, air density ρ_a and wind velocity (in m/s)

$$u_a = 5 + (\sin(2\pi t/T) - 3)\sin(2\pi x_1/L_{x_1})\sin(\pi x_2/L_{x_2})\,,$$
$$v_a = 5 + (\sin(2\pi t/T) - 3)\sin(2\pi x_2/L_{x_2})\sin(\pi x_1/L_{x_1})\,,$$

with $T = 4$ days. Initially, the ice is 2 m thick and the ice concentration increases linearly from 0 in the west to 1 in the east, so that the mean ice thickness h varies from 0 to 2 m. The mean wind pushes the ice into the northeast corner where it gradually piles up until it becomes sufficiently thick to be stopped. We will use $\Delta_{min} = 2 \cdot 10^{-9}\ \mathrm{s}^{-1}$ (Hibler III, 1979).

5.2. Convergence of B- and C-grid discretizations of the mEVP method

We start with an examination of convergence and stability of the mEVP scheme on B- and C-grids. It suffices to consider the first external time level

(Kimmritz et al., 2015). Recall the C1 and C2 discretizations of the shear viscosities at nodal points. In the C1 case, the nodal shear viscosity is the average of the shear viscosities defined at adjacent cells; in the C2 case, it is computed with fewest possible averages of the contributing variables. Figure 4 plots the residuals:

$$\left(\sum_{i,j} \frac{\alpha^2 |\boldsymbol{\sigma}_{ij}^{p+1} - \boldsymbol{\sigma}_{ij}^{p}|^2}{\alpha^2 |\boldsymbol{\sigma}_{ij}^{2} - \boldsymbol{\sigma}_{ij}^{1}|^2} + \sum_{i,j} \frac{\beta^2 |\mathbf{u}_{ij}^{p+1} - \mathbf{u}_{ij}^{p}|^2}{\beta^2 |\mathbf{u}_{ij}^{2} - \mathbf{u}_{ij}^{1}|^2} \right)^{1/2},$$

of the subcycling at the first time level for B-, C1- and C2-grid discretizations. We weighted the single contributions in the definition of the residual by the inverse of the first residuals of the subcycling in order to put each of the contributions on equal footing. Convergence within numerical working precision is

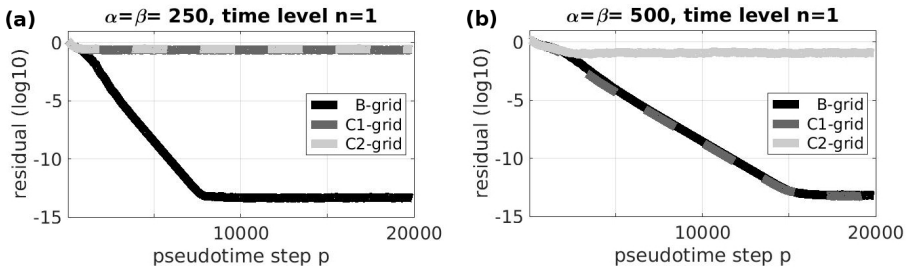


Figure 4: Residuals of the first time level of the full nonlinear problem for the B-, the C1- and the C2-grid discretization, and for different choices of $\alpha = \beta$ ((a) $\alpha = \beta = 250$, (b) $\alpha = \beta = 500$).

reached for $\alpha = \beta = 250$ after $0.75 \cdot 10^4$ subcycling steps only in the B-grid case, and for $\alpha = \beta = 500$ after $1.5 \cdot 10^4$ subcycling steps for the B-grid discretization and the C1-grid case. The C2-grid discretization does not converge in any case. (Note that Lemieux and Tremblay (2009) also needed $O(10^4)$ nonlinear steps in their Picard iteration.) We cannot give a rigorous explanation for this behavior, but we hypothesize that the viscosity computation in the C2-case prevents the discrete analogue of (2) to be satisfied exactly. For the remaining schemes we recover the expected behavior (see also Kimmritz et al., 2015): higher values of α and β guarantee stability but slow down the speed of convergence. In agreement with our analysis above, the stability constraints appear to be stricter for

the C1-discretization than for the B-grid discretization. However, if the C1-grid scheme converges, its convergence rate is only marginally slower than the convergence rate of the B-grid scheme.

5.3. Convergence of B- and C-grid discretizations of the aEVP method

In Figure 5 we compare the convergence rates of the aEVP approach with α and β computed by (12) to the mEVP scheme (Bouillon et al., 2013, Kimmritz et al., 2015) with fixed $\alpha = \beta = 500$. The parameters for the aEVP scheme are set to $c = (0.01\pi)^2$, $\tilde{c} = 4$, and $(\alpha, \beta) \geq 5$. Note, that we set c to a very small value. This implies, that we deal with scales that are two orders of magnitude larger than the grid scale and thus consider basin scale. It can only reflect the fact that within the first time step there is still no detail in the velocity field and thus allows us to use this small value. On later time levels we expect a larger variety of scales in the velocity field, which requires larger values for c.

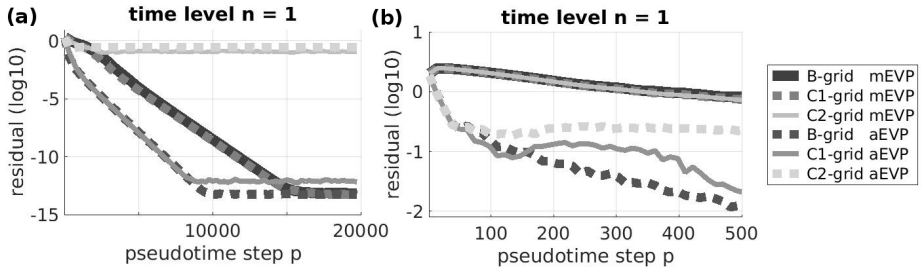


Figure 5: Residuals in the subcycling on the first time level for different discretizations. Graph (a) plots the entire convergence behavior, graph (b) is a zoom into the first 500 subcycling steps.

As in the mEVP case, the B-grid and the C1-grid discretizations lead to convergence, but the C2-case does not converge. Convergence in the C2-case is also not gained for different settings of c and $\tilde{c}$ (not shown). The convergence of the adaptive approach for the B- and the C1-grid case is faster than for mEVP by a factor of 3, but the final residual for the C1-grid is slightly larger than for the mEVP scheme. As in practice the affordable number of subcycling steps is probably 500 or less (Kimmritz et al., 2015), we concentrate on the residual

development over the first 500 sybcycling steps in Figure 5 (b). Compared to the mEVP approach we see a reduction in the residual of more than one order of magnitude in the convergent cases. Even for the C2-case the residuals are smaller for aEVP. In agreement with our theoretical analysis, there are more oscillations in the residuals for the C1-grid case than for the B-grid case.

Errors may accumulate over finite time intervals. We simulate the ice evolution over one month with $N = 500$ subcycling steps and examine the performance of the aEVP scheme implemented now in the MITgcm with the C1-grid arrangement. Because of the oscillatory decrease of the residuals at the first time level we use a larger stabilizing parameter $c = (0.5\pi)^2$. In the beginning of the subcycling at time level 1440 the residuals in the momentum and in the stress equations of the aEVP scheme are almost an order of magnitude smaller than the ones of the mEVP scheme with $\alpha = \beta = 500$ (Figure 6 (a)). The residuals of the momentum equations in both schemes decrease at a similar rate in both schemes. In the subcycling of the mEVP scheme the residuals of the stress equations converge with a rate, which is similar to the rate of the momentum equations. The residual of the stress equations in the aEVP scheme increases in the first 20 subcycling steps, which might be explained by the adaptation of the α field to the updated fields on the new time level. After this 'initial' phase the residual in the stress equations decreases at an increased rate, such that at the end of the subcycling the residual of the stress equations in the aEVP sheme is about 1.5 orders smaller than the residual of the stress equations in the mEVP scheme. At the end of the subcycling at time level 1440, α (and thus β) is very small ($\alpha = 5$) in the large region of weak ice (Figure 6 (b)). Kimmritz et al. (2015) demonstrated that the number of subcycling steps to reach full convergence for the given example is of the order of $N_{EVP} = 40\alpha$. Thus, we can presume, that the scheme reached full convergence in those regions with $N = 500$ subcycling steps. However, since α^{-1} and β^{-1} define the pseudotime step in units of Δt, too small values may lead to a loss of accuracy of the pseudotime iterations. Thus we recommend to always impose lower bounds for α and β.

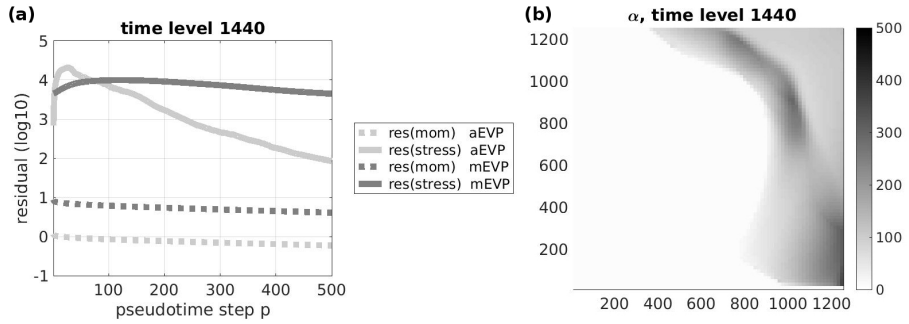


Figure 6: (a) Residual development for the subcycling at time level 1440 of the aEVP scheme with $c = (0.5\pi)^2$ and $\tilde{c} = 4$, and the mEVP scheme with $\alpha = \beta = 500$. The residuals in the momentum equations (res(mom)) are given by $(\sum_{ij} \beta_{ij}^2 |\boldsymbol{u}^{p+1} - \boldsymbol{u}^p|^2)^{1/2}$, the residuals in the stress equations (res(stress)) are computed as $(\sum_{ij} \alpha_{ij}^2 |\boldsymbol{\sigma}^{p+1} - \boldsymbol{\sigma}^p|^2)^{1/2}$. (b) The α field at the end of the subcycling at time level 1440 of the aEVP scheme with 500 subcycling steps.

Beside sufficient accuracy, the aEVP scheme should guarantee smoothness of the solution. According to Kimmritz et al. (2015), the corresponding mEVP scheme with $\alpha = \beta = 250$ shows noise in the divergence field. Figure 6 indicates that in the aEVP scheme large values of α are only used in a small region in the lower right corner of the domain where the ice is strong. Outside this region, α ranges between 200 and 300 over the area with ice concentrations between 0.8 and 1.

To evaluate the aEVP scheme we use a converged VP solution determined with the JFNK solver of the MITgcm (Losch et al., 2014) with a C1-grid discretization and a residual reduction of order 10^{-9} in each time step as reference solution, and also consider solutions of the mEVP scheme with $\alpha = \beta = 500$ to illustrate the improvements through adaptivity. We note that the solutions of the mEVP scheme with $\alpha = \beta = 500$ and $N = 20000$ (full convergence) coincide with the solutions determined with the JFNK solver, but N as large as this would be too expensive for practical applications (climate simulations). To examine the effect of the lower bounds of α and β in the adaptive scheme with $N = 500$, 300 and 200 subcycling steps, we explore the cases $(\alpha, \beta) \geq 5$ and $(\alpha, \beta) \geq 50$. In Figure 7 we present the deviations in the divergence field from

the reference solution after one month of integration. We note, that the results in the Δ field, the ice concentration and ice thickness are of similar quality (not shown). The aEVP and mEVP schemes have been run with $N = 500, 300$ and 200 subcycling steps (columns from left to right). The black lines in the graphs mark the boundary with ice concentration of 0.01. The regions left of them correspond to open water. The errors seen there are of little relevance and will not be discussed.

Compared to the mEVP solution with $N = 500$ subcycling steps, any of the aEVP solutions leads to a remarkable reduction in the errors of the adaptive scheme even for the case of $N = 200$ subcycling steps. According to Figure 7 the aEVP scheme shows virtually no errors in the area covered with ice for $N = 500$. The errors increase only slightly for $N = 300$ and even for the case of $N = 200$ they remain small and are much smaller than the errors for mEVP.

For $N = 200$, the residuals of the aEVP scheme in regions with strong ice show noisy behavior for the lower bound for α and β of 5 (graphs (a) – (c)). This noise vanishes when we increase the lower bound to 50 (graphs (d) – (f)). We relate the emergence of noise in the first case to an excessively large pseudotime step and hence reduced pseudotime iteration accuracy. These errors accumulate already in the early stage of the simulation. A lower bound substantially larger than 50, however, is not advisable as it may have adverse effects on the convergence in large parts of the ice covered regions thus jeopardizing the benefits of the aEVP scheme.

6. Conclusion and Outlook

The present work has two main results: First, the modified EVP scheme (Bouillon et al., 2013) is less stable on a C-grid, than on a B-grid, and convergence of the scheme on a C-grid is sensitive to the implementation of the viscosities. Second, we introduced the new adaptive EVP scheme, which locally respects stability constraints as derived in (Kimmritz et al., 2015), and shows improved convergence properties while guaranteeing stability in regions with

higher stability constraints.

The main advantage of the mEVP implementation (Bouillon et al., 2013) of the commonly used viscous-plastic rheology over the traditional EVP implementation (Hunke and Dukowicz, 1997) is the decoupling of the parameters of the subcycling from the external time stepping. The mEVP is formulated as a pseudotime solver of ice dynamics with the VP rheology. Convergent solutions can only be obtained if the iterative process is numerically stable. In this paper we elucidated the sensitivity of the convergence of the mEVP approach to the detail of numerical discretization. An elementary eigenvalue analysis revealed that the mEVP implementation on a B-grid is more stable than on a C-grid. If both schemes are stable and converge, their convergence rates are comparable. The convergence on C-grids, however, is sensitive to the implementation of the viscosities. We considered two versions of implementation that have been suggested in literature; one of them (C2) does not converge to the VP solution and is always contaminated by noise, while the other (C1) does so under stable conditions. The lack of convergence for the C2 implementation might be related to its lack of energy consistency (Bouillon et al., 2013). A rigorous explanation for this behavior is still missing, but we hope that this result on its own provides an important message to modellers.

In our earlier work we showed that, on the one hand the mEVP parameters α and β need to be sufficiently large to ensure stability. They define the frequency of the numerical oscillations. The requirement $\gamma/(\alpha\beta) \ll 1$ limits the frequency of these oscillations to sufficiently low values to be well represented by the pseudotime iterations. On the other hand, large values of α and β necessitate a large number of subcycling steps to reach convergence, which makes the scheme very expensive for practical applications (long climate simulations). Emphasizing the dependence of γ on the mesh resolution we pointed out that the tendency to use finer meshes in large-scale ocean modelling implies larger values of γ, hence larger values for α, β and N. This would increase the computational cost of sea ice codes further. This argument is valid for any change in the model parameters that effects an increase in γ.

The main point of the present study is the new adaptive implementation of the mEVP approach. Instead of being constant, the parameters α and β are locally adjusted at each pseudotime step (12). The (constant) number of iterations N is selected experimentally so as to provide reasonable accuracy everywhere in the ice covered domain.

By choosing α and β adaptively we guarantee global stability. Since the adaptive α and β are relatively low in wide areas of the ice covered domain, convergence in those regions is improved with respect to the mEVP method. Our test experiments reveal a substantial error reduction in the aEVP solutions compared to the mEVP solutions even for smaller N. This is a big gain in terms of computational costs. In preliminary tests, 500 subcycling steps already raised the cost of the sea ice component to about 50% of the ocean model, which is undesirably large. In a next step, the aEVP approach has to be applied to a realistic scenario in order to test the overall performance and to learn about admissible N. This will be the subject of a companion paper.

The aEVP approach can be especially useful for models that are based on locally refined meshes, as it guarantees stability in the most refined areas. It will also lead to advantages in areas where the ice is weak or of relatively low concentration by reducing α and β and hence improving convergence there.

The new adaptive approach can be further augmented in several ways. The version described here still contains parameters that have to be selected experimentally, yet they can be estimated at run time. For instance, the factor c can be assessed through the local smoothness of the velocity field. The other question is the optimal choice of N based on the information of the distribution of α and β. While it is difficult to change N during the subcycling, it is possible to select different N at different external time steps. These opportunities will be explored in future work.

In closing, we like to point out that there are other recently published sea ice rheologies that also involve elasticity, such as the elastic plastic anisotropic rheology (Tsamados et al., 2013) or the elasto brittle approach (Girard et al., 2009, 2011, Bouillon and Rampal, 2015). If the appropriate schemes are solved

explicitly through pseudotime stepping, a stability analysis similar to ours or to Kimmritz et al. (2015) may serve as a basis for designing an approach similar to the aEVP.

S. Bouillon and P. Rampal. Presentation of the dynamical core of neXtSIM, a new sea ice model. *Ocean Modelling*, 91:23–37, 2015. doi: http://dx.doi.org/10.1016/j.ocemod.2015.04.005.

S. Bouillon, M. A. Morales Maqueda, V. Legat, and T. Fichefet. An elastic-viscous-plastic sea ice model formulated on arakawa b and c grids. *Ocean Modelling*, 27:174–184, 2009.

S. Bouillon, T. Fichefet, V. Legat, and G. Madec. The elastic-viscous-plastic method revisited. *Ocean Modelling*, 71:2–12, 2013.

L. Girard, J. Weiss, J. M. Molines, B. Barnier, and S. Bouillon. Evaluation of high-resolution sea ice models on the basis of statistical and scaling properties of Arctic sea ice drift and deformation. *Journal of Geophysical Research: Oceans*, 114(C8), 2009.

L. Girard, S. Bouillon, J. Weiss, D. Amitrano, T. Fichefet, and V. Legat. A new modeling framework for sea-ice mechanics based on elasto-brittle rheology. *Annals of Glaciology*, 52(57):123–132, 2011.

W. D. Hibler III. A Dynamic Thermodynamic Sea Ice Model. *J. Phys. Oceanogr.*, 9:815–846, 1979.

W. D. Hibler III and C. F. Ip. The effect of sea ice rheology on Arctic buoy drift. *J. P.Dempsey and Y. D. S.Rajapakse (eds.) ASME AMD*, 207, Ice Mechanics: 255–263, 1995.

E.C. Hunke. Viscous-plastic sea ice dynamics with the EVP model: Linearization issues. *J. Comp. Phys.*, 170:18–38, 2001.

E.C. Hunke and J.K. Dukowicz. An Elastic-Viscous-Plastic model for sea ice dynamics. *J. Phys. Oceanogr.*, 27:1849–1867, 1997.

E.C. Hunke and W.H. Lipscomb. CICE: The Los Alamos sea ice model documentation and software users manual. Tech. Rep. *T-3 Fluid Dynamics Group, Los Alamos National Laboratory, Los Alamos NM 87545*, 2008.

M. Kimmritz, S. Danilov, and M. Losch. On the convergence of the modified elastic-viscous-plastic method for solving the sea ice momentum equation. *J. Comp. Phys.*, 296:90–100, 2015.

J.-F. Lemieux and B. Tremblay. Numerical convergence of viscousplastic sea ice models. *J. Geophys. Res.*, 114, C05009, 2009.

J.-F. Lemieux, B. Tremblay, J. Sedláček, P. Tupper, St. Thomas, D. Huard, and J.-P. Auclair. Improving the numerical convergence of viscous-plastic sea ice models with the jacobian-free newtonkrylov method. *J. Comp. Phys.*, 229(8): 2840–2852, 2010. doi: http://dx.doi.org/10.1016/j.jcp.2009.12.011.

J.-F. Lemieux, D. Knoll, B. Tremblay, D.M. Holland, and M. Losch. A comparison of the Jacobian-free Newton-Krylov method and the EVP model for solving the sea ice momentum equation with a viscous-plastic formulation: a serial algorithm study. *J. Comp. Phys.*, 231(17):5926–5944, 2012.

M. Losch and S. Danilov. On solving the momentum equations of dynamic sea ice models with implicit solvers and the elasticviscousplastic technique. *Ocean Modelling*, 41:42–52, 2012.

M. Losch, D. Menemenlis, J.-M. Campin, P. Heimbach, and C. Hill. On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*, 3(1-2):129–144, 2010.

M. Losch, A. Fuchs, J.-F. Lemieux, and A. Vanselow. A parallel Jacobian-free NewtonKrylov solver for a coupled sea ice-ocean model. *J. Comp. Phys.*, 257, Part A(0):901–911, 2014.

M. Tsamados, D. L. Feltham, and A. V. Wilchinsky. Impact of a new anisotropic rheology on simulations of arctic sea ice. *J. Geophys. Res.: Oceans*, 118(1): 91–107, 2013. doi: 10.1029/2012JC007990.

J. Zhang and W.D Hibler. On an efficient numerical method for modeling sea ice dynamics. *J. Geophys. Res.*, 102:8691–8702, 1997.

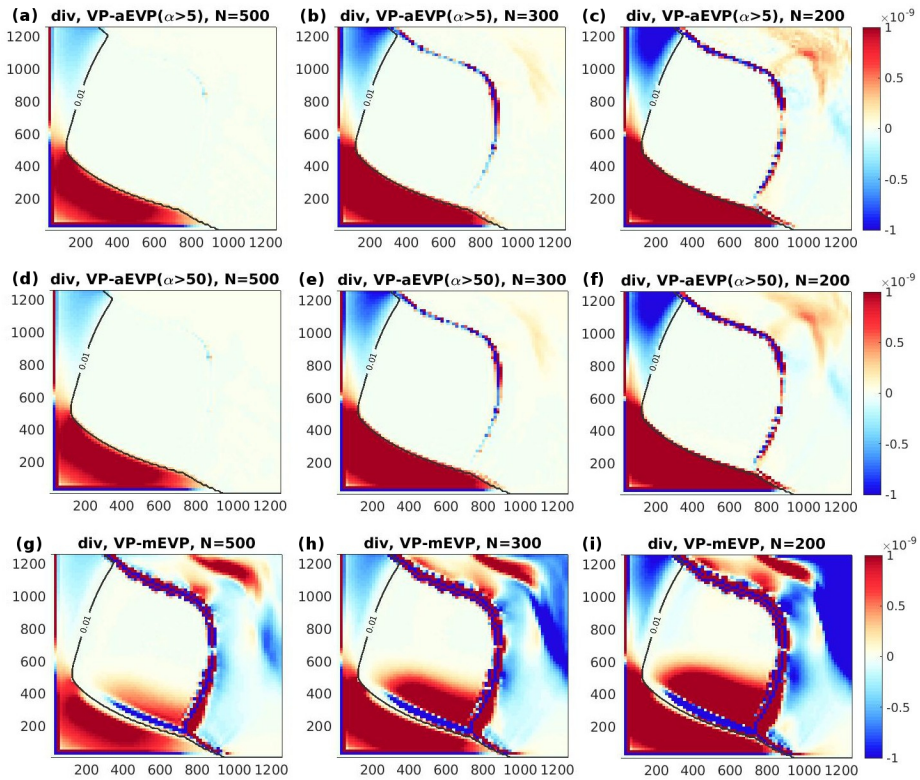


Figure 7: Differences in the divergence field between the reference solution and the aEVP solution with the lower bound $(\alpha, \beta) > 5$ for (a) $N = 500$, (b) $N = 300$ and (c) $N = 200$. Graphs (d)–(f): Same as (a)–(c) with the lower bound $(\alpha, \beta) > 50$. Graphs (g)–(i): Differences in the divergence field between the reference solution and the mEVP solution with $\alpha = \beta = 500$ and (g) $N = 500$, (h) $N = 300$ and (i) $N = 200$. The black lines are the isolines of ice concentration for $a = 0.01$. All of these runs use the C1-grid formulation.