Journal of Computational Physics **148,** 579–604 (1999)
Article ID jcph.1998.6130, available online at http://www.idealibrary.com on IDEAL®

Implementation of a Barotropic–Baroclinic Time Splitting for Isopycnic Coordinate Ocean Modeling

Robert L. Higdon

Department of Mathematics, Oregon State University, Corvallis, Oregon 97331-4605
E-mail: higdon@math.orst.edu

Received October 15, 1997; revised November 2, 1998

The motions seen in numerical models of ocean circulation typically include rapidly moving external gravity waves, which are essentially two-dimensional, and a variety of other motions which are much slower and generally three-dimensional. In a number of ocean models, the computational problems resulting from the multiple time scales are addressed by splitting the fast and slow dynamics into separate subproblems that are solved by different techniques. The present paper addresses the practical implementation of such a splitting for the case of isopycnic ocean modeling, in which the vertical coordinate is density or some other related quantity. During the implementation, a major task is to develop formulas suitable for usage with rapidly varying bottom topography. Compared to an earlier splitting, the present splitting has improved stability properties when analyzed in a simple linearized setting. When tested in the same model for which the earlier splitting was developed, the revised splitting substantially reduces a numerically induced sloshing pattern that is seen in the model. This phenomenon appears to be related to a residual term that appears in the momentum equations. In some examples, the residual is one to two orders of magnitude smaller in the case of the revised splitting.

Key Words: ocean circulation; numerical solution of partial differential equations; stability analysis; time discretization; barotropic-baroclinic time splitting.

1. INTRODUCTION

Numerical models of ocean circulation typically admit motions varying on a wide range of time scales. The motions include rapidly moving external gravity waves, whose dynamics are similar to those found in the two-dimensional shallow-water equations that describe the motions of a hydrostatic fluid of constant density. The remaining motions, such as currents and interval gravity waves, have velocities that can be at least two orders of magnitude smaller. In order to deal with the computational difficulties imposed by the disparity in

0021-9991/99 $30.00

time scales, several ocean circulation models (e.g., [1, 3, 10, 14]) model the fast motions with a two-dimensional barotropic system of equations and the remaining slow motions with a three-dimensional baroclinic system. The baroclinic system is solved explicitly with a long time step that is appropriate for resolving the slow motions, while the barotropic system is solved by separate techniques. A more extensive literature survey is given in the introduction to [7]. The present paper addresses the splitting process in the context of isopycnal coordinates, for which the vertical coordinate is density or some other related quantity. In this type of model, a vertical discretization divides the fluid into layers that are approximately immiscible, so water masses are tracked automatically due to the choice of coordinate system. Other advantages of isopycnal coordinates are discussed in [1, 7].

Perhaps the first barotropic–baroclinic splitting for isopycnal models was the one developed by Bleck and Smith [1]. Higdon and Bennett [7] analyzed this splitting for the case of a simple linearized model having a level bottom and two fluid layers, and they found that the method in [1] can yield unstable computational algorithms due to inexactness in the splitting. The equations that model the slow motions actually admit some energy moving on the fast scale, so the Courant–Friedrichs–Lewy condition is violated for those equations. Sources of inexactness include the decompositions of the velocity and pressure fields and the derivation of the barotropic momentum equation. The latter is a prognostic equation for the vertical average of the horizontal velocity, and in this equation the pressure-forcing term is equivalent to the one from the shallow water equations. Higdon and de Szoeke [8] subsequently found that the pressure term is the main source of instability. In an isopycnal model, the horizontal pressure forcing is provided by the gradient of the Montgomery potential $M = \alpha p + gz$. If the pressure term in the barotropic system in [1] is replaced by the vertical average of ∇M, then the instability is essentially eliminated in the linearized setting analyzed in [7, 8]. Additional improvements are given by some time stepping schemes developed by Hallberg [5]. Essentially, in the splitting developed in [1], the barotropic momentum equation neglects the variations in density over the depth of the fluid, whereas the splitting in [8] incorporates the vertical structure more explicitly.

The present paper extends the preceding work, as follows. One goal is to develop a method for implementing the splitting of [8] in a nonlinear isopycnic model with strongly varying bottom topography. A second goal is to test this method when applied to the Miami Isopycnic Coordinate Ocean Model (MICOM) [1, 2], which is the model for which the splitting in [1] was originally developed.

When pursuing the first goal, a particular concern is developing a suitable representation of $\overline{\nabla M}$, the vertical average of ∇M. Two methods for representing this average are discussed in Section 3. In one approach, which is termed a bottom–up method, one first determines M_b, the Montgomery potential in the bottom layer. One then represents ∇M in all other layers in terms of ∇M_b prior to the vertical averaging. However, at locations where layer interfaces intersect the bottom of the fluid domain, representations of $\overline{\nabla M}$ based on this approach can include terms that contain discontinuities that are not already intrinsic to the problem. This point is discussed in Section 3.1.1. These discontinuities act as sources of velocity convergence which can generate spurious gravity waves in the computed solution. In some numerical experiments using such an approach, erratic behavior appeared quickly, followed by an ultimate blowup that apparently resulted from nonlinear interactions in the system. An alternate approach, which is termed a top–down method, is developed in Sections 3.1.2 and 3.2. With this method the idea is to represent M_1, the Montgomery potential in the top layer, in terms of changes in layer thicknesses and then represent ∇M

in all other layers in terms of ∇M_1. The quantity ∇M_1 is continuous as long as the free surface is smooth, and with this approach the final representation of $\overline{\nabla M}$ does not contain discontinuous terms of the type encountered with the bottom–up approach.

In the following discussions of splittings, the term "original splitting" will refer to the splitting developed by Bleck and Smith [1], and the term "revised splitting" will refer to the splitting developed by Higdon and de Szoeke [8]. In the context of a nonlinear model with varying bottom topography, the revised splitting is implemented using the top–down method mentioned above, as opposed to the bottom–up method. However, in the discussion of linearized stability given in Section 2.3, this distinction is immaterial, due to the simplicity of the case considered there. The concept of top–down versus bottom–up also does not apply to the original splitting in [1], as it does not employ a vertical average of the Montgomery gradient.

The above splittings were compared in numerical computations involving a regional configuration of MICOM. The spatial resolution used in these particular tests is adequate for showing some of the main features of oceanic circulation, but it is not fine enough to give high resolution of boundary currents or to resolve the eddies that are shed from currents. In these experiments, the revised and original splittings produced similar results, in the sense that snapshots of various flow fields yielded results that are similar, although not identical. However, the original splitting produced a substantial sloshing pattern that appears to be a numerical artifact, instead of a realistic physical phenomenon. The sloshing can be seen in the net lateral transport of fluid and in the depth-dependent circulation. The sloshing was reduced substantially when the revised splitting was used. It was initially anticipated that the improved stability of the revised splitting would make it possible to run the model with less time smoothing than with the original splitting. Time smoothing was originally introduced into that model in order to suppress the computational mode that is allowed by the leapfrog scheme that is used to solve the baroclinic equations. However, the effect of this mode is sufficiently strong for reasons related to the implementation of the mixed layer that it did not appear possible to run the revised splitting with substantially reduced smoothing in the present series of tests.

An outline of the paper is the following. Section 2 sketches the revised and original splittings, and Section 3 describes a method for implementing the revised splitting. The results of the numerical experiments are described in Section 4. A summary and some conclusions are given in Section 5.

2. OUTLINE OF THE SPLITTINGS

This section describes the revised splitting that will be implemented and tested in later sections. An outline of the earlier splitting of Bleck and Smith [1] is also given here.

2.1. *Summary of the Splitting Process*

The following discussion is based on the primitive equations as applied to a vertically discrete, layered model. Assume that the fluid consists of R layers with constant specific volumes $\alpha_1, \ldots, \alpha_R$, with index 1 referring to the uppermost layer and index R referring to the bottom layer. (The specific volume is the reciprocal of density.) Let $p_r(x, y, t)$ denote the pressure at the top of layer r; the pressure at the free surface is assumed to be $p_1 = 0$. The elevation and pressure at the bottom of the fluid will be denoted by $z_b(x, y)$ and

$p_b(x, y, t) = p_{R+1}(x, y, t)$, respectively. Let $\Delta p_r = p_{r+1} - p_r$ denote the pressure difference across layer r, and let $\Delta\alpha_r = \alpha_r - \alpha_{r+1}$ denote the increment in α across the interface at the bottom of layer r. Let $M_r(x, y, t) = \alpha_r p + gz$ denote the Montgomery potential in layer r; a calculation involving the hydrostatic condition $\partial p/\partial z = -\alpha^{-1} g$ shows that M is independent of depth in a layer for which α is independent of depth. Finally, let $\mathbf{u}_r(x, y, t) = (u_r, v_r)$ denote the horizontal velocity in layer r. A simplified version of the primitive equations in then

$$\frac{\partial \mathbf{u}_r}{\partial t} + (\mathbf{u}_r \cdot \nabla)\mathbf{u}_r + f\mathbf{k} \times \mathbf{u}_r = -\nabla M_r \tag{2.1a}$$

$$M_r = M_{r+1} + p_{r+1}\Delta\alpha_r \tag{2.1b}$$

$$\frac{\partial \Delta p_r}{\partial t} + \nabla \cdot (\mathbf{u}_r \Delta p_r) = 0. \tag{2.1c}$$

(Also see [1, 7].) Here, $\nabla = (\partial/\partial x, \partial/\partial y)$, and $f\mathbf{k} \times \mathbf{u}_r$ denotes the Coriolis terms $(-f v_r, f u_r)$.

The jump condition (2.1b) can be derived by observing that in the quantity $M = \alpha p + gz$, α is discontinuous at an interface, whereas p and z are continuous. Equation (2.1c) describes the conservation of mass, as $g^{-1}\Delta p_r$ is the mass per unit horizontal area in layer r. A realistic ocean model would also incorporate physical effects such as thermodynamics, viscosity, surface forcing, and mixing between layers, but the system (2.1) is sufficient for illustrating the essential ideas behind a barotropic–baroclinic time splitting.

For the momentum equation in the barotropic system, we use a vertical average of (2.1a). Define a mass-weighted, vertically averaged velocity by

$$\bar{\mathbf{u}}(x, y, t) = \sum_{r=1}^{R} \frac{\Delta p_r}{p_b} \mathbf{u}_r. \tag{2.2}$$

The quantity $g^{-1} p_b \bar{\mathbf{u}}$ is the net rate of horizontal mass transport, and for a linearized problem the formula (2.2) approximates the projection of the velocity field onto the external mode [7]. The vertical average of (2.1a) can be written in the form

$$\frac{\partial \bar{\mathbf{u}}}{\partial t} + f\mathbf{k} \times \bar{\mathbf{u}} = -\overline{\nabla M} + \mathbf{G}, \tag{2.3}$$

where $\overline{\nabla M}$ is defined in analogy to (2.2). The residual term $\mathbf{G}(x, y, t)$ represents the net effect of those quantities whose vertical averages are not represented explicitly in (2.3). In the present simplified case, $\mathbf{G}$ consists of the vertical average of the nonlinear terms and a term involving the time derivatives of the weight coefficients $\Delta p_r / p_b$. In a realistic ocean model, $\mathbf{G}$ would also include such quantities as stress and viscosity. The implementation of $\mathbf{G}$ is discussed below.

A baroclinic momentum equation can be obtained by subtracting (2.3) from (2.1a) to obtain a prognostic equation for the baroclinic velocity $\mathbf{u}'_r = \mathbf{u}_r - \bar{\mathbf{u}}$. A comparison with (2.2) shows that $\mathbf{u}'$ has mass-weighted vertical average zero.

For a decomposition of the pressure field, we use a splitting that was introduced by Bleck and Smith [1]. Let $p'_b(x, y)$ denote the pressure at the bottom of the fluid when the system is in some reference state, such as an initial state or static state. For any (x, y, t) let $\eta(x, y, t)$ denote the relative perturbation in bottom pressure from that state, so that

$p_b(x, y, t) = (1 + \eta) p_b'$. For the sake of usage in Section 3.1, also assume $M_1 = g z_{\text{top}}$ is constant in the reference state. Given the pressure field p and the preceding definition of η, define a depth-dependent quantity p' so that $p = (1 + \eta) p'$ throughout the fluid. The pressure increments across each layer then satisfy

$$\Delta p_r(x, y, t) = (1 + \eta(x, y, t)) \Delta p_r'(x, y, t). \tag{2.4}$$

The quantity η is intended to represent the effects of external gravity waves, and p' is expected to vary only on a slow time scale. Typically $|\eta| \ll 1$. The decomposition (2.4) then induces a splitting of the mass equation (2.1c); see [1] or [7].

The decomposition (2.4) is based on the idea that an external wave causes all fluid layers to thicken or thin by approximately the same proportion. Such a statement is not exact. The decomposition $\mathbf{u}_r = \mathbf{u}_r' + \bar{\mathbf{u}}$ of the velocity field is also inexact, to a similar degree. However, the analysis of Higdon and de Szoeke [8] indicates that these errors do not have a significant impact on the stability of the coupled barotropic–baroclinic algorithm.

The barotropic momentum equation of Bleck and Smith [1] has the form

$$\frac{\partial \bar{\mathbf{u}}}{\partial t} + f \mathbf{k} \times \bar{\mathbf{u}} = -\alpha_0 \nabla(p_b' \eta) + \mathbf{G}^*, \tag{2.5}$$

where α_0 is a reference value of specific volume. In [1], the residual term is denoted by $\bar{\mathbf{u}}_t^*$ instead of $\mathbf{G}^*$. The quantity $p_b' \eta$ is the perturbation in bottom pressure from the reference state. For the case of a hydrostatic fluid of constant density, $p_b' \eta$ is also the pressure perturbation induced by changes in the elevation of the free surface. The term $-\alpha_0 \nabla(p_b' \eta)$ is thus equivalent to the pressure forcing that appears in the shallow water equations for a hydrostatic fluid of constant density $1/\alpha_0$.

With the barotropic momentum equations (2.3) and (2.5), the residual terms $\mathbf{G}$ and $\mathbf{G}^*$ can be implemented as follows. (Also see [1, 8].) Suppose that the solution is to be advanced from time t_n to time $t_{n+1} = t_n + \Delta t$, where Δt is the (long) baroclinic timestep. First, advance the baroclinic momentum equation with the residual term deleted. The condition that $\mathbf{u}'$ has vertical mean zero is then used to adjust $\mathbf{u}'$ and obtain the residual. The computed $\mathbf{G}$ (or $\mathbf{G}^*$) is transferred to the barotropic momentum equation, where it serves as a forcing term when the barotropic equations are solved on the time interval $[t_n, t_{n+1}]$. In [1] and in the computations described in Section 4, the barotropic equations are solved explicitly with short substeps. In this computation, the term $\mathbf{G}$ (or $\mathbf{G}^*$) can vary with (x, y) but is treated as constant in t for $t_n < t < t_{n+1}$.

In the barotropic momentum equation (2.5), the pressure forcing term neglects the vertical variations of density within the fluid, whereas the alternative equation (2.3) incorporates the vertical structure explicitly. In the case of (2.3), the residual term $\mathbf{G}$ represents the net effect of various quantities discussed earlier. In the case of (2.5), the residual term contains these effects, plus the difference between $\alpha_0 \nabla(p_b' \eta)$ and $\overline{\nabla M}$. An explicit representation of the latter is developed in Section 3.1, and from that representation it can be seen that $\alpha_0 \nabla(p_b' \eta) - \overline{\nabla M}$ includes terms varying on the fast barotropic time scale. Using $-\overline{\nabla M}$ in place of $-\alpha_0 \nabla(p_b' \eta)$ has the effect of pulling this difference out of the residual and treating it with an explicit formula. In some computational examples discussed in Section 4, this step reduces the size of the residual by one to two orders of magnitude.

2.2. *Stability*

The effect of replacing $-\alpha_0 \nabla(p_b' \eta)$ with $-\overline{\nabla M}$ can also be seen in a linear stability analysis. Higdon and Bennett [7] analyzed the stability of the splitting of Bleck and Smith [1] for the special case of a linearized system with one horizontal dimension, two layers, and flat bottom topography. Higdon and de Szoeke [8] subsequently analyzed the stability of the revised splitting in the same setting and also for the case of two horizontal dimensions. In [7] the parameter α_0 was taken to be the value of α in the lower layer, as this choice simplified some formulas in the analysis. However, in the current version of MICOM, which is discussed in Section 4, the value of α_0 is chosen to be 1 cm^3/g for convenience. This value is greater than all of the values of α that are used in the layers in that model. Because of the discrepancy between the analysis and the model, we briefly revisit the analysis and test the dependence of stability on the value of α_0.

In the analysis in [7, 8], it is assumed that the system is discretized with respect to space by using second-order centered finite differences on a staggered grid. The barotropic equations are solved exactly with respect to t in order that time discretization errors for those equations do not mask the effects of the splitting or the baroclinic timestepping. In order to analyze stability, we perform a Fourier transform with respect to x and examine the evolution of the system through time. Equivalently, consider special solutions of the form $\lambda^n e^{ikx} q$, where λ is a complex scalar, n is the time index, k is a real wavenumber, and q is a vector. In general, λ and q depend on k. Because of periodicity, one can assume $-\pi \le k\Delta x \le \pi$, where Δx is the space step. An algorithm would be stable if, for each wavenumber, $|\lambda| \le 1$ for all λ and if linear independence enables all solutions to be written as linear combinations of solutions of the form $\lambda^n e^{ikx} q$.

For the sake of brevity, we do not give the details of the analysis, but instead show plots of $|\lambda|$ versus $k\Delta x$ for a sample set of model parameters. For these plots it is assumed that $(\alpha_1 - \alpha_2)/\alpha_2 = 0.004$, where α_1 and α_2 are the specific volumes of the upper and lower layers, respectively. It is also assumed that the pressure difference across the upper layer is one-quarter the bottom pressure. The baroclinic timestep Δt is chosen so that the baroclinic Courant number is $c_1 \Delta t/\Delta x = 0.4$, where c_1 is the internal wave speed. The leap frog method is used to discretize the baroclinic equations with respect to t. As shown in [8], the calculation of λ can be reduced to an eigenvalue problem having six nonzero eigenvalues.

In Fig. 1, the upper graph shows a plot of $|\lambda|$ versus $k\Delta x$ for the splitting of Bleck and Smith, with $\alpha_0 = \alpha_2$. For this plot, values of $k\Delta x$ were sampled in increments of $\pi/2000$. The absolute values of all eigenvalues are shown in the plot, so multiple curves are present. For most k there exists λ for which $|\lambda| > 1$, so the method is unstable in the present setting.

The lower frame in Fig. 1 shows plots of $|\lambda|$ for the case $\alpha_0 = \alpha_1$, which is the value of α in the upper layer. The instability remains, although the maximum value of $|\lambda|$ is slightly smaller. One difference with the preceding case is that there is a stronger prevalence of modes for which $|\lambda| < 1$; for those particular modes, the algorithm is dissipative. Some additional plots, not shown here, indicate that this pattern persists as α_0 is increased to values greater than those actually found in the fluid. As α_0 increases, the maximum value of $|\lambda|$ decreases gradually but remains greater than 1. The minimum values drop further below 1, which indicates greater dissipation in some modes.

Figure 2 shows a plot of $|\lambda|$ for the revised splitting. The model parameters are the same as in Fig. 1. In the calculation leading to this plot, the baroclinic quantities appearing in $\overline{\nabla M}$ are interpolated linearly with respect to t between consecutive baroclinic time levels. With

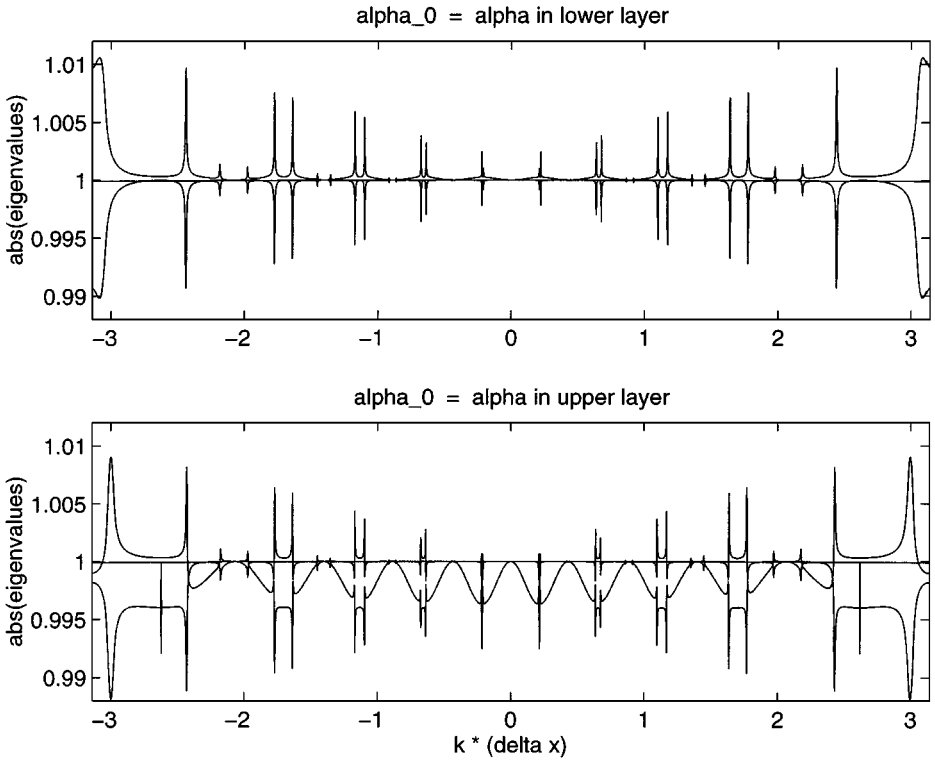

FIG. 1. Plots of $|\lambda|$ versus $k\Delta x$ for the original splitting developed in [1]. In this splitting, the pressure gradient term $-\alpha_0\nabla(p_b'\eta)$ is used in the barotropic momentum equation. The present analysis considers a linearized system with two fluid layers and one horizontal dimension. In the top frame, it is assumed that α_0 is equal to the value of α in the lower layer; for the bottom frame, α_0 is the value of α in the upper layer. Model parameters are specified in the text.

the revised splitting, the eigenvalues have absolute value exactly equal to 1, except for a few spikes defined on some extremely narrow intervals. The spikes can be eliminated by using a time stepping scheme of Hallberg [5], which employs a prediction step and a correction step for both the barotropic and baroclinic equations. In computational experiments, it has not

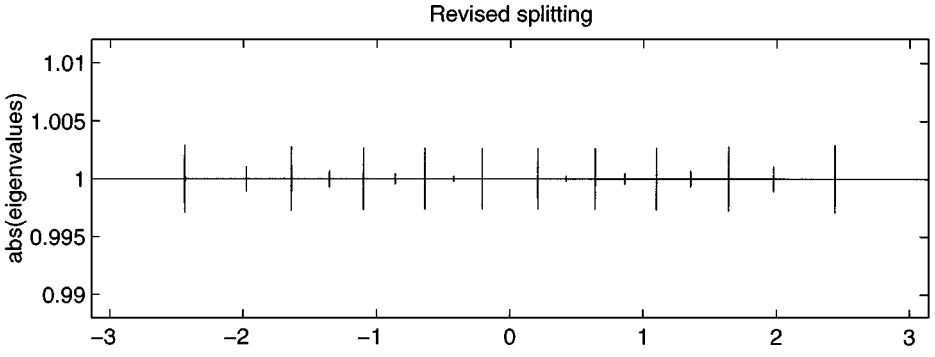

FIG. 2. Plot of $|\lambda|$ versus $k\Delta x$ for the revised splitting implemented in the present paper. In this case, the pressure forcing term in the barotropic momentum equation is $-\overline{\nabla M}$.

been evident that the spikes have a significant effect on the computation, given the effects of viscosity and other factors that are present in an ocean circulation model.

3. IMPLEMENTATION OF THE REVISED SPLITTING

In this section we describe some aspects of implementing the splitting of [8] in a nonlinear model with variable bottom topography. One of the main steps is to express $\overline{\nabla M}$ in terms of the prognostic variables in the system.

3.1. *Representation of* $\overline{\nabla M}$ *in the Horizontally Continuous Case*

The basic approach taken here is to represent M_1, the Montgomery potential in the top layer, in terms of changes in layer thicknesses. The interface condition (2.1b) then yields M in all other layers, and a representation for $\overline{\nabla M}$ can be obtained.

In the following discussion, it is assumed that all fluid layers exist at all horizontal positions in the ocean. However, the interfaces between layers may intersect the bottom or top boundaries of the fluid, so some of the layers can have (essentially) zero thickness in some regions. This technique of massless layers is used, for example, by Bleck *et al.* [1, 2] and Oberhuber [12] in isopycnic ocean models and by Hsu and Arakawa [9] in an isentropic model of the atmosphere.

3.1.1. Some difficulties related to $\overline{\nabla M}$. Before giving the derivation of $\overline{\nabla M}$ sketched above, it would be useful to discuss another approach that in some respects might seem more natural. At the bottom of the fluid, the Montgomery potential satisfies the boundary condition $M_b = \alpha_b p_b + g z_b$ which is easily evaluated once the layer thicknesses are known. The interface condition can then be used to relate M in all other layers to M_b. A corresponding representation of $\overline{\nabla M}$ would then involve ∇M_b. The calculation of the latter can be facilitated by using the pressure splitting $p = (1+\eta)\, p'$ to obtain $\nabla M_b = \nabla M'_b + \nabla(\alpha_b p'_b \eta)$, where $M'_b = \alpha_b p'_b + g z_b$ denotes the Montgomery potential in the bottom layer at the reference state that is used to define the pressure splitting. The quantity $\alpha_b p'_b \eta$ is several orders of magnitude smaller than $\alpha_b p'_b$, so separating this term from the remainder of M_b can yield a more accurate calculation of the effect of its gradient. In addition, $\nabla M'_b$ is independent of time and need not be updated during the computation.

The preceding approach was used to obtain the representation of $\overline{\nabla M}$ given in Eq. (2.12) of Higdon and de Szoeke [8]. This representation was adequate for the numerical computations described in that paper, which assumed level bottom topography. However, in later computations involving highly irregular bottom topography, this representation produced erratic behavior and an ultimate blowup. Following are some observations that may help to explain the irregular behavior and motivate the alternate approach to $\overline{\nabla M}$ that is developed in the present paper.

For the moment, consider a two-layer fluid bounded below by topography that varies linearly and suppose that the reference state is a stationary state for which the free surface and fluid interface are level. Also assume that the free surface is located at $z = 0$. At this state, the Montgomery potential in the upper layer is $M'_1 = \alpha_1 p'_1 + g z_{\text{top}} = g z_{\text{top}} = 0$, and the Montgomery potential in the lower layer is $M'_b = M'_1 - p'_2 \Delta\alpha = -p'_2 \Delta\alpha$. In a region where the lower layer has positive thickness, p'_2 is constant, and $\nabla M'_b = 0$. On the other hand, in a region where the lower layer has zero thickness, p'_2 varies linearly, and $\nabla M'_b$ is

a nonzero constant. The quantity $\nabla M_b'$ then has a jump discontinuity at any point where the interface intersects the bottom topography. A similar analysis shows that analogous conclusions can be reached for a multilayered fluid at any point where a smooth interface intersects bottom topography at a nontangential angle. This situation is a consequence of using massless layers at the bottom of the fluid.

However, $\overline{\nabla M}$ is itself continuous, assuming that the free surface and bottom topography are described by continuously differentiable functions and that the top of each layer is continuously differentiable on any region where the layer has positive thickness. To see this, observe that these hypotheses, together with the relation $M_r = gz_{\text{top}} - \sum_{j=2}^{r} p_j \Delta\alpha_{j-1}$, imply that ∇M_r is continuous in layer r, except perhaps for jump discontinuities at locations where the top of layer r intersects the bottom topography. However, in the definition of $\overline{\nabla M}$ the quantity ∇M_r is multiplied by the weighting coefficient $\Delta p_r / p_b = \Delta p_r' / p_b'$, which tends to zero at intersection points. The product of these quantities is continuous. The quantity $\overline{\nabla M}$ is thus continuous, although it could have discontinuous first derivatives.

Now suppose that one is using a representation of $\overline{\nabla M}$ in which $\nabla M_b'$ appears as a separate term. Taken in isolation, this term would act as a forcing term on the barotropic flow, and the discontinuities in this forcing would provide localized sources of velocity convergence which could generate spurious gravity waves in the solution. However, since $\overline{\nabla M}$ is continuous, the discontinuities in $\nabla M_b'$ must be cancelled by jumps of opposite sign appearing in the remaining terms in $\overline{\nabla M}$, and the spurious wave are not actually generated. This observation depends on an exact cancellation of discontinuities appearing in different terms. If these terms are then approximated by finite differences, this cancellation may not necessarily occur. This appears to be the source of the erratic behavior in some numerical experiments involving some implementations of the representation of $\overline{\nabla M}$ derived in [8]. The ultimate blowup may have been due to nonlinear interactions involving the spurious waves.

In the preceding discussion, it was assumed that the top of each layer is smooth on any region where the layer has positive thickness. However, the top of a layer is not smooth at a location where it intersects the free surface at a nontangential angle. In that case, an analysis similar to the above shows that $\overline{\nabla M}$ could be discontinuous. However, this situation has not caused noticeable difficulties in numerical experiments. In this case, the discontinuity in $\overline{\nabla M}$ appears to be a dynamic configuration to which the system can adjust. For example, consider a two-layer fluid having a level free surface at $z = 0$, and assume that the thickness of the upper layer is zero for $x < x_0$ and proportional to $x - x_0$ for $x > x_0$. In this case, $M_1 = M_2 = \overline{\nabla M} = 0$ for $x < x_0$. On the other hand, if $x > x_0$ then $M_1 = 0$, $\partial M_2/\partial x < 0$, and the x-component of $-\overline{\nabla M}$ is positive. Since $-\overline{\nabla M}$ provides forcing to $\partial \bar{\mathbf{u}}/\partial t$, the discontinuity in $-\overline{\nabla M}$ provides a source of barotropic fluid divergence at $x = x_0$. More generally, if the intersection of the interface with the free surface is smoothed slightly, then $-\overline{\nabla M}$ undergoes a rapid transition near $x = x_0$, and again a source of barotropic fluid divergence is obtained. Such a source induces a drop in the free surface elevation which serves to moderate the divergence; in this sense, the system adjusts to the jump in $\overline{\nabla M}$. On the other hand, a discontinuity in $\nabla M_b'$ represents a static forcing that acts forever, and in the experiments cited above the system appeared unable to adjust to this kind of forcing in a reasonable manner.

It should be noted that there are practical limitations to how smoothly $\overline{\nabla M}$ can vary when the system is discretized in the horizontal in the presence of strongly varying bottom topography. In that case, at a fixed time, there can be different numbers of nonzero layers

at different grid points, and the bottom topography can then be visualized as a stairstep pattern. For example, in the model configuration used in Section 4, the number of nonzero layers can jump by several units between adjacent grid points, as illustrated in Fig. 3. In such situations, $\overline{\nabla M}$ could contain a jump when the number of nonzero layers changes, due to a change in the number and character of the terms being averaged. For example, consider a two-layer model that admits one nonzero layer at all grid points for which $x < x_0$ and two nonzero layers at all points where $x > x_0$. Suppose that the fluid has a level free surface at $z = 0$, the interface between layers varies linearly for $x > x_0$, and the thickness of the lower layer approaches a positive value as $x \to x_0^+$. In this case $M_1 = 0$ for all x, and M_2 varies linearly for $x > x_0$. It follows that $\overline{\nabla M}$ is zero for $x < x_0$ but has nonzero values immediately to the right of x_0. More generally, a jump in $\overline{\nabla M}$ is possible if there are n_1 nonzero layers for $x < x_0$ and n_2 nonzero layers for $x > x_0$, with $n_2 - n_1 > 1$; if, for some configuration of the system, $\overline{\nabla M}$ does not have a jump, then another configuration with a jump can be constructed by varying the shapes of the tops of layers $n_1 + 1, \ldots, n_2$ for x slightly greater than x_0. The precise behavior of $\overline{\nabla M}$ at $x = x_0$ would depend on the structure of the grid and the method of discretizing $\overline{\nabla M}$. In any case, abrupt variations of $\overline{\nabla M}$ are possible in neighborhoods of points where the number of nonzero layers changes. This situation appears intrinsic to the problem and not an artifact of any particular formula for $\overline{\nabla M}$.

With the formulation of $\overline{\nabla M}$ that is discussed above and used in [8], one would have to balance carefully the discretization of different terms in order to avoid jumps that are greater than those that are intrinsic to the problem. On the other hand, the difficulty involving $\nabla M_b'$ is related to the massless layers at the bottom of the fluid, and an alternative is to abandon the use of such layers. However, this would entail difficulties related to tracking the moving boundaries of nonzero layers.

The approach taken in the present paper is to sidestep these difficulties by formulating $\overline{\nabla M}$ differently. Essentially, the preceding method is a bottom–up approach which relates the Montgomery potential in each layer to the Montgomery potential in the lowest layer. Instead, we use here a top–down approach that begins with a representation of the Montgomery potential in the top layer.

3.1.2. A formula for $\overline{\nabla M}$. The Montgomery potential in the top layer is $M_1(x, y, t) = g z_{\text{top}}(x, y, t)$, where z_{top} is the perturbation in the elevation of the free surface relative to the mean elevation $z = 0$. Let $\Delta z_r(x, y, t)$ denote the thickness of layer r, and let $\Delta z_{r,\text{ref}}(x, y)$ denote the thickness of layer r when the system is at the reference state used to define the pressure splitting $p = (1 + \eta) p'$. In Section 2.1 it was assumed that, in this state, the free surface is level. It follows that $M_1 = g \sum_{r=1}^{R} (\Delta z_r - \Delta z_{r,\text{ref}})$. The hydrostatic condition $\partial p / \partial z = -\alpha^{-1} g$ then yields

$$M_1 = \sum_{r=1}^{R} \alpha_r (\Delta p_r - \Delta p'_{r,\text{ref}}),$$

where $\Delta p'_{r,\text{ref}}(x, y)$ denotes the pressure difference across layer r at the reference state. Now insert the pressure decomposition (2.4) to obtain

$$M_1(x, y, t) = \bar{\alpha} p_b' \eta + \sum_{r=1}^{R} \alpha_r (\Delta p_r' - \Delta p'_{r,\text{ref}}), \tag{3.1a}$$

where

$$\bar{\alpha}(x, y, t) = \sum_{r=1}^{R} \alpha_r \frac{\Delta p'_r}{p'_b} \tag{3.1b}$$

denotes the mass-weighted vertical average of α. In (3.1a), the quantity $p'_b\eta$ varies on the fast time scale, and the other quantities vary on the slow time scale. The summation in (3.1a) represents some effects of internal waves; the terms in this sum cancel partially, due to varying signs in the vertical direction.

The Montgomery potential in an arbitrary layer r can then be written

$$\begin{aligned} M_r &= M_1 - p_2\Delta\alpha_1 - \cdots - p_r\Delta\alpha_{r-1} \\ &= M_1 - (1+\eta)N_r, \end{aligned}$$

where $N_1 = 0$ and $N_r = p'_2\Delta\alpha_1 + \cdots + p'_r\Delta\alpha_{r-1}$ for $r > 1$. The quantity N resembles Montgomery potential, except that it increases downward with increments of the form $p'\Delta\alpha$. The vertical average of ∇M is then

$$\begin{aligned} \overline{\nabla M} &= \sum_{r=1}^{R} \frac{\Delta p'_r}{p'_b} \nabla M_r \\ &= \nabla M_1 - \sum_{r=1}^{R} \frac{\Delta p'_r}{p'_b} \nabla \left[p'_b(1+\eta)\frac{N_r}{p'_b} \right], \end{aligned} \tag{3.2}$$

where ∇M_1 is calculated from (3.1). The quantity N_r/p'_b has units of α and satisfies $0 < N_r/p'_b < \Delta\alpha$, where $\Delta\alpha = \alpha_1 - \alpha_R$ denotes the variation in α over the total depth of the fluid. The quantity $p'_b(1+\eta)$ is equal to the total bottom pressure p_b.

The barotropic momentum equation of Bleck and Smith [1] uses the gradient term $-\alpha_0\nabla(p'_b\eta)$, where α_0 is a reference value of α. A comparison with (3.1) and (3.2) shows that the difference between $-\alpha_0\nabla(p'_b\eta)$ and $-\overline{\nabla M}$ consists of the gradient of the summation term in (3.1a) plus terms of order $\Delta\alpha$. This difference includes terms that vary on the fast barotropic time scale.

3.2. *Representation of $\overline{\nabla M}$ in the Horizontally Discrete Case*

For definiteness, assume that the horizontal discretization of the governing equations is based on the C-grid of Arakawa [6, 11]. This discretization is used in the model tested in Section 4. The quantities M, p, η, and N are defined at "mass points" in the centers of grid cells, and the values of normal components of velocity are defined at the centers of edges of grid cells. Let u and v denote the components of velocity with respect to x and y, respectively. If the center of a grid cell is located with integer indices (i, j), then denote the u-points associated with that cell with indices (i, j) and $(i+1, j)$. Similarly, denote the v-points with indices (i, j) and $(i, j+1)$.

The following discussion considers centered differences of M_r with respect to x that are taken to be approximations to $\partial M_r/\partial x$ at u-points. A similar discussion applies to differences with respect to y that represent values at v-points.

When computing vertical averages of discretizations of $\partial M_r/\partial x$, it is necessary to use weight coefficients $\Delta p'_r/p'_b$ evaluated at u-points. For a fixed value of time, let $\Delta p'_r(u, i, j)$

denote an approximation to $\Delta p_r'$ at the u-point with indices (i, j). The calculation of this quantity is discussed in Section 3.3. The quantity $p_b'(u, i, j) = \sum_{r=1}^{R} \Delta p_r'(u, i, j)$ is an approximation to p_b' at u-point (i, j), and the weight coefficient $\Delta p_r'(u, i, j)/p_b'(u, i, j)$ can be used when calculating vertical averages at that point. However, such coefficients are not used when calculating $\bar{\alpha}$ in the expression (3.1) for M_1. All of the quantities in (3.1) are located at mass points, and in the calculation of $\bar{\alpha}$ at mass point (i, j) it suffices to use the values of $\Delta p_r'$ and p_b' that are already defined at that point.

In analogy to (3.2), a discretization of the x-component of $\overline{\nabla M}$ is given by

$$
\begin{aligned}
(\overline{\nabla M})_x &\approx \frac{M_1(i, j) - M_1(i-1, j)}{\Delta x} \\
&\quad - \sum_{r=1}^{R} \frac{\Delta p_r'(u, i, j)}{p_b'(u, i, j)} \left(\frac{p_b(i, j)\frac{N_r}{p_b'}(i, j) - p_b(i-1, j)\frac{N_r}{p_b'}(i-1, j)}{\Delta x} \right) \\
&= \frac{M_1(i, j) - M_1(i-1, j)}{\Delta x} - \frac{A(u, i, j)p_b(i, j) - B(u, i, j)p_b(i-1, j)}{\Delta x}, \qquad (3.3)
\end{aligned}
$$

where

$$
\begin{aligned}
A(u, i, j) &= \sum_{r=1}^{R} \frac{\Delta p_r'(u, i, j)}{p_b'(u, i, j)} \frac{N_r(i, j)}{p_b'(i, j)} \\
B(u, i, j) &= \sum_{r=1}^{R} \frac{\Delta p_r'(u, i, j)}{p_b'(u, i, j)} \frac{N_r(i-1, j)}{p_b'(i-1, j)},
\end{aligned} \qquad (3.4)
$$

M_1 is given by the expression in (3.1), and $p_b = p_b'(1 + \eta)$. The quantities $A(u, i, j)$ and $B(u, i, j)$ vary on the slow time scale, and p_b varies on the fast time scale.

If the barotropic equations are solved explicitly by using short substeps of a baroclinic time interval $[t_n, t_{n+1}]$, then the quantities in $\overline{\nabla M}$ that vary on the slow time scale can be interpolated in t between times t_n and t_{n+1}. The barotropic quantities are updated at each short timestep.

3.3. *Weighting Coefficients*

The calculation of $\overline{\nabla M}$ requires values of the weighting coefficients $\Delta p_r'/p_b'$ at velocity points. However, these coefficients are naturally defined at mass points, so an interpolation between mass points is required. During this interpolation, a problem that must be avoided is the creation of false pressure gradients in regions of varying bottom topography. For an example of this situation, suppose that the system is at a static state characterized by a horizontal free surface, horizontal interfaces between layers, $\eta = 0$, and zero forcing. Consider a u-point with indices (i, j), and let $p_r'(i, j)$ and $p_r'(i-1, j)$ denote the baroclinic pressures at the top of layer r at the adjacent mass points. Also suppose that the top of layer r rests on the sea floor at mass point $(i-1, j)$ but not at mass point (i, j), and the top of that layer is higher at mass point $(i-1, j)$ than it is at mass point (i, j). This situation would occur if the top of layer r intersects bottom topography between these mass points, and layer r is then continued with zero layer thickness for decreasing x. The interface condition (2.1b), together with $M_1 = 0$, then implies $M_r(i-1, j) > M_r(i, j)$. If the weighting coefficient

$\Delta p'_r(u, i, j)/p'_b(u, i, j)$ is nonzero, then the vertical average of the discrete $\partial M/\partial x$ contains a nonzero contribution from layer r. Since this vertical average provides forcing to $\overline{u}$, it is possible that the system will start moving, even though it should remain stationary. The scheme for calculating weighting coefficients must therefore yield $\Delta p'_r(u, i, j) = 0$ in this situation. A simple average of $\Delta p'_r(i-1, j)$ and $\Delta p'_r(i, j)$ would not have this property.

Some weighting coefficients are computed in MICOM in order to calculate vertical averages of velocities. These coefficients satisfy the requirements stated above, so this weighting strategy is given here as an example. Let

$$p'_r(u, i, j) = \min\left(p'_b(i-1, j),\, p'_b(i, j),\, \frac{p'_r(i-1, j) + p'_r(i, j)}{2}\right).$$

This quantity provides an approximation to the interface pressure at the u-point (i, j). Then let

$$\Delta p'_r(u, i, j) = p'_{r+1}(u, i, j) - p'_r(u, i, j)$$

to obtain a pressure increment across layer r. The bottom pressure is then given by

$$p'_b(u, i, j) = \sum_{r=1}^{R} \Delta p'_r(u, i, j) = \min(p'_b(i-1, j),\, p'_b(i, j)).$$

In regions that are located away from the bottom of the fluid, $\Delta p'_r(u, i, j)$ is simply the average of $\Delta p'_r(i, j)$ and $\Delta p'_r(i-1, j)$. However, in the situation described in the preceding paragraph, $p'_r(u, i, j) = p'_b(i-1, j)$ and $p'_{r+1}(u, i, j) = p'_b(i-1, j)$, so $\Delta p'_r(u, i, j) = 0$ as desired. A comparison with (3.4) and (3.5) shows that $\overline{\nabla M} = 0$ for this case, so false pressure gradients are avoided.

3.4. *Other Issues of Implementation*

The following are some other aspects of implementing the splitting that is tested in Section 4. These issues are mostly specific to MICOM, although some of the motivations for time smoothing apply more generally.

3.4.1. Pressure forcing in the upper layer. In MICOM the uppermost layer represents the mixed layer. In the physical ocean the mixed layer is vertically homogeneous due to mixing caused by wind forcing and thermal convection. In the model, the density of the uppermost layer is independent of depth but can vary with (x, y, t). Because of this fact, the horizontal pressure forcing in layer 1 is $\nabla M_1 - p\nabla\alpha_1$, instead of ∇M_1. To derive this result, observe that the hydrostatic condition $\partial p/\partial z = -\alpha^{-1}g$, together with the assumption of zero atmospheric pressure, implies that the pressure in layer 1 is $p(x, y, z, t) = \alpha_1^{-1}(x, y, t)g(z_{\text{top}}(x, y, t) - z) = \alpha_1^{-1}(M_1 - gz)$. The horizontal pressure forcing term is then $\alpha_1\nabla p = \nabla M_1 - \alpha_1^{-1}(M_1 - gz)\nabla\alpha_1 = \nabla M_1 - p\nabla\alpha_1$.

In the computations described in Section 4, this quantity is approximated by $\nabla M_1 - p'\nabla\alpha_1$ for convenience. The quantity p' varies with depth, so an intermediate value must be chosen when implementing this term. In addition, values of p' are needed at velocity points, so an interpolation between adjacent mass points is needed. These steps are accomplished by

using half an harmonic mean. For example, in the approximation to $\partial M_1/\partial x - p'\partial\alpha_1/\partial x$ at u-point (i, j), the value of p' is taken to be $p_2'(i-1, j)p_2'(i, j)/(p_2'(i-1, j) + p_2'(i, j))$. Here, $p_2'(i-1, j)$ and $p_2'(i, j)$ are the baroclinic pressures at the bottom of layer 1 at the mass points adjacent to u-point (i, j).

Strictly speaking, the above issues should also be taken into account when constructing the vertically averaged barotropic momentum equation. That is, $\overline{\nabla M}$ should be replaced by $\overline{\nabla M} - (\Delta p_1'/p_b')p'\nabla\alpha_1$. However, in some numerical experiments this modification did not appear to produce significant changes in the computed results. For the sake of simplicity, $-\overline{\nabla M}$ is used as the forcing term for the barotropic momentum equation in the computations described in Section 4. The forcing for the baroclinic momentum equations is thus $-(\nabla M_1 - p'\nabla\alpha_1 - \overline{\nabla M})$ for layer 1 and $-(\nabla M_r - \overline{\nabla M})$ for all other layers.

3.4.2. Time smoothing. In general, a smoothing operation may be needed when implementing the barotropic–baroclinic splitting described above. As noted earlier, in a discrete model $\overline{\nabla M}$ can contain jumps due to abrupt changes in the number of nonzero layers over which the average is taken, and these jumps could generate spurious gravity waves. One would like to suppress such waves, and some time smoothing can accomplish this. It should also be noted that the idea of a barotropic–baroclinic splitting is based on the decomposition of the solution into external and internal modes. Strictly speaking, this decomposition applies to linearized models with level bottom topography for which analytical solutions can be constructed by the method of separation of variables. (See [7].) However, in a realistic ocean model a barotropic–baroclinic splitting is fundamentally inexact. This leaves the possibility of erratic behavior that would need to be removed from the computation.

In MICOM, there is an additional reason for time smoothing that is independent of the issues discussed elsewhere in this paper. The baroclinic equations are presently solved with the leapfrog time stepping scheme, which admits a computational mode consisting of grid-scale oscillations with respect to time. In addition, the treatment of the mixed layer in this model allows for sudden detrainment of large amounts of water from the mixed layer into the interior layers. This detrainment provides an impulsive force that can stimulate the computational mode. If left unchecked, these oscillations can ultimately cause a failure in the computation, so they need to be removed by a filtering device.

In standard MICOM, this filter has the form $\bar{f}^n = 0.125\,\bar{f}^{n-1} + 0.75\,f^n + 0.125\,f^{n+1}$, where f^n denotes the value of a generic field f at timestep n, and $\bar{f}^n$ denotes the filtered value of the field at that time. This filter is used to modify the solution at time step n after the solution at timestep $n+1$ has been computed. (Also see Haltiner and Williams [6].) The temperature and salinity equations use the same filter, but with weighting coefficients (1/64, 31/32, 1/64).

In the tests of the revised splitting described in Section 4, a different strategy is used. The baroclinic equations are solved with a filter having coefficients (0.01, 0.98, 0.01). On the other hand, the solution of the barotropic equations involves the following time-averaging of barotropic variables. With the parameters used in these particular computations, the barotropic equations are solved explicitly with 40 barotropic timesteps per baroclinic timestep. We continue this computation out to step 50 and compute time-averages of $\bar{\mathbf{u}}$ and $p_b'\eta$ over steps 30 through 50. These time-averages are communicated to the baroclinic equations and are used as the starting point for the next barotropic integration.

This averaging is used for the following reason. The leapfrog scheme is an example of the midpoint rule for numerical integration. When applied to the generic equation $u'(t) = F(t)$,

this rule implies

$$u(t_{n+1}) = u(t_{n-1}) + \int_{t_{n-1}}^{t_{n+1}} F(\tau)\, d\tau$$
$$\approx u(t_{n-1}) + 2\Delta t F(t_n),$$

which is a reasonable formula when F is slowly varying. However, when F is rapidly varying, the sampled value $F(t_n)$ is sensitive to small phase shifts in F, and in that case it would be better to replace $F(t_n)$ with a time average. The time-averaging employed in Section 4 uses data from approximately one quarter of the interval $[t_{n-1}, t_{n+1}]$.

4. NUMERICAL COMPUTATIONS

The preceding strategy was implemented in MICOM in order to compare the new and original splittings. For these computations, the model domain is a region of the north and equatorial Atlantic Ocean having northern and southern boundaries at latitudes approximately 65° N and 27° S, respectively. This region of the Atlantic is embedded in a rectangular region that is discretized in the horizontal with 128 grid intervals in both the east–west and north–south directions, which gives a resolution of approximately 0.9° at the equator. A masking operation is used to delete land masses from the computation. Solid-wall boundary conditions are imposed at the northern and southern boundaries.

Vertical discretization is accomplished by dividing the fluid into 16 layers. The uppermost layer represents the mixed layer; in this model the density of this layer can vary with time and horizontal position but not with vertical position. The other layers in the model have constant density. The specific volume of the top layer is 0.97589 cm^3/g at the initial time. The other layers in the model have specific volumes 0.97530, 0.97472, 0.97423, 0.97382, 0.97348, 0.97320, 0.97297, 0.97278, 0.97262, 0.97248, 0.97236, 0.97226, 0.97218, 0.97212, and 0.97208.

In Section 3.1.1 it was mentioned that the intersection of layer interfaces with the bottom of the fluid domain can cause difficulties with implementing $\overline{\nabla M}$. This situation is illustrated in Fig. 3 which shows a portion of the model domain consisting of the Caribbean Sea and the Gulf of Mexico. The figure contains a collection of integers, each of which is located at a mass point on the grid. For each point, the integer specifies the number of nonzero fluid layers at that point at the initial time. The figure shows that the number of nonzero layers can change dramatically from one grid point to the next. This strong variation in bottom topography provides an illustration of the difficulties discussed in Section 3.1.1. In numerical computations of the type described below, the bottom–up representation of $\overline{\nabla M}$, which is criticized in Section 3.1.1, yielded irregular behavior and an eventual blowup. Details of the computational results with this representation are not given here. On the other hand, the top–down representation of $\overline{\nabla M}$ yielded reasonable results, as described below.

In these computations, the free surface elevation and horizontal velocity components are set to zero at the initial time. The temperature field and layer thicknesses are initialized to zonal (east–west) averages of climatological data. Evolution of the model through time is then determined by forcing from climatological values of winds, precipitation, evaporation, atmospheric moisture, atmospheric temperature, and solar radiation. Some of these fields are specified at monthly intervals, and others are specified quarterly. These data sets are

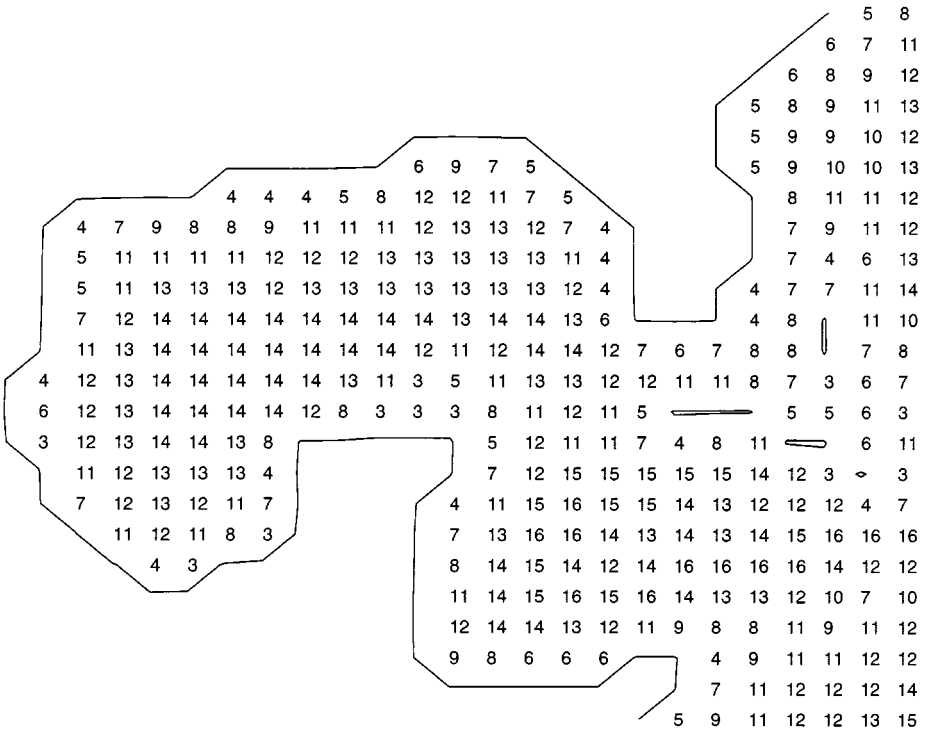

FIG. 3. Illustration of variable bottom topography. The figure shows a portion of the model domain consisting of the Caribbean Sea and the Gulf of Mexico. The integers in the plot are located at mass points on the spatial grid. For each point, the integer indicates the number of nonzero fluid layers at that point at the initial time.

then interpolated with respect to time in order to provide forcing to the model. The seasonal cycle in the forcing produces a seasonal cycle in the computed solution.

In this set of computations the time step used for the baroclinic equations is 1.5 h, or 1/16 day. The barotropic equations are solved explicitly by using the forward–backward method with 40 barotropic timesteps per baroclinic step. The numerical computations primarily include tests of the following versions of the model:

(a) A modified version of MICOM which incorporates the top–down formulation of the revised splitting as described in Sections 3.1.2 and 3.2.

(b) The original, unmodified MICOM. In the forcing term $-\alpha_0 \nabla(p_b' \eta)$ in the barotropic momentum equation, the parameter α_0 was chosen for convenience to be 1 cm^3/g during the development of that model.

(c) A modified version of MICOM that uses the original splitting as in (b), but with the same time-smoothing scheme used for (a). A comparison of (a) and (b) is of considerable interest, but it leaves open the question of how the different time-smoothing schemes in (a) and (b) might affect the results. A comparison of (a) and (c) isolates the effect of the splitting itself.

Computations were performed on a 32-node partition of a Connection Machine CM5. Execution times for the revised model (a) were roughly 6% to 9% longer than for the original model (b). Part of the extra time is due to the vertical summations required to implement $\overline{\nabla M}$. This cost is intrinsic to the method described here. The remainder of the extra time

is due to the additional computation of barotropic substeps as described in Section 3.4.2. This cost is a consequence of the time-averaging used in these particular tests.

Each version of the model was integrated from day 0 to day 3600. In this model 3600 days are regarded as 10 years, as each month is assumed to be 30 days long during the specification of the forcing functions. Snapshots were made of various fields at day 3600 and at day 3420 (9.5 years) from the output from versions (a), (b), and (c). These fields include surface velocity, layer thicknesses, salinity, and mixed layer density. Snapshots of the same fields with different versions of the model are generally similar, although not identical.

Some examples are given in Figs. 4 and 5 which show the interfaces between fluid layers. Figure 4 shows two meridional (north–south) cross sections that were produced with the revised splitting for day 3600. Sections are shown at longitudes 44.40° W and 36.30° W. The horizontal coordinate is the latitude in degrees, with the northernmost latitude to the left, and the vertical coordinate is the depth in meters. The shaded regions indicate the sea

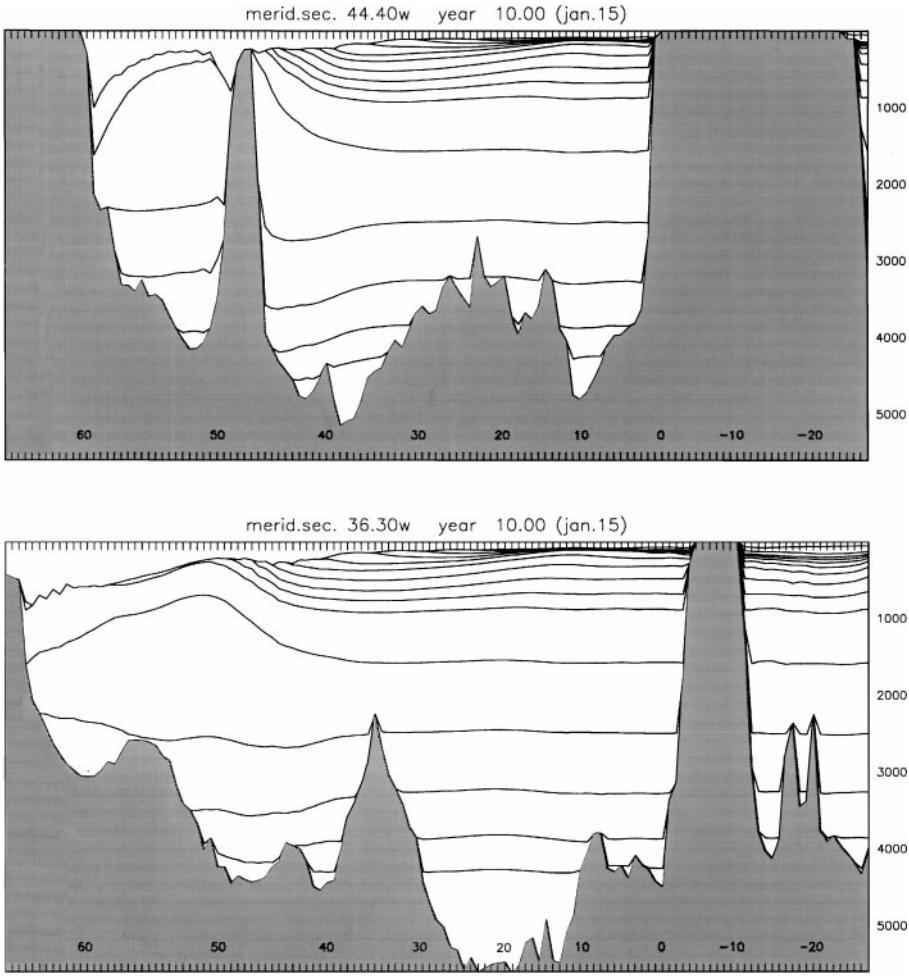

FIG. 4. Vertical, north–south cross sections at day 3600 of the solution computed with the revised splitting, as implemented in version (a) of the model. The horizontal coordinate is the latitude in degrees, and the vertical coordinate is the depth in meters. The shaded regions indicate the sea floor and land masses. The curves in the interior of the fluid region show the interfaces between layers.

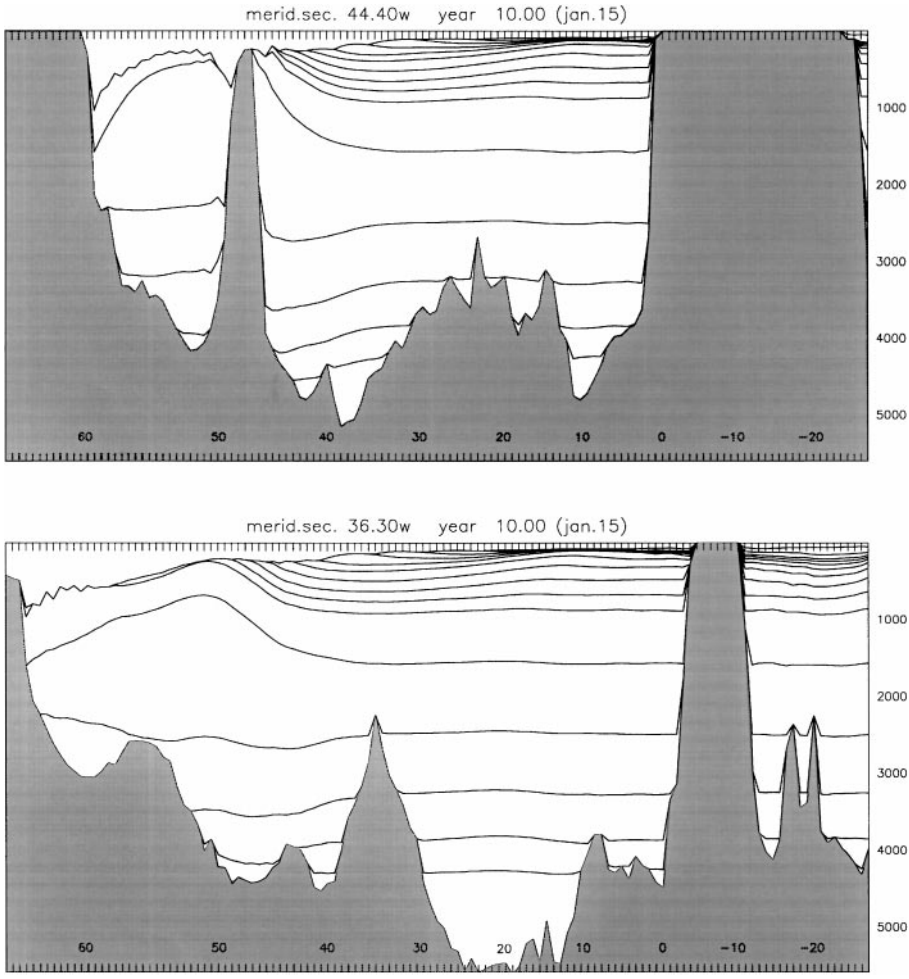

FIG. 5. Cross sections produced with the original splitting, as implemented in version (c).

floor and land masses. The layer interfaces are illustrated by the curves that are approximately horizontal. The uppermost layer represents the mixed layer; this layer is relatively thick in the northernmost regions, which is typical of wintertime conditions. Several of the interior layers outcrop into the mixed layer in the latitude range 40° N to 45° N, which is approximately the location of the Gulf Stream. Figure 5 shows the corresponding cross sections that were obtained with the original splitting, as implemented in version (c) of the model. Figures 4 and 5 are generally very similar. As mentioned in Section 3.1, an issue encountered during the implementation of $\overline{\nabla M}$ is the effect of interfaces intersecting the bottom topography. Figure 4 does not indicate any problems in this regard, with the present implementation.

As noted in Section 3.4.2, substantial time smoothing is needed in this model because of the computational mode that is allowed by the timestepping scheme that is used to solve the baroclinic equations. The computational mode is independent of the issue of barotropic–baroclinic splitting. It appears that the smoothing blurs the differences in stability between the two splittings tested here, and this helps to account for the general similarity of the results obtained with versions (a), (b), and (c).

However, the results are not identical, and some differences can be seen in some diagnostics of meridional mass flux that were taken at regular intervals throughout the computation. On a discrete set of latitudes, at each diagnostic time, the meridional mass transport in each layer was summed east–west in order to obtain a layer-by-layer measure of the net north–south transport. A vertical sum of these transports then gives the total meridional transport as a function of latitude. The layer-by-layer transport can also be used to produce a meridional streamfunction that illustrates the depth-dependent circulation. As discussed below, each of these diagnostics shows a sloshing pattern that is substantially reduced by using the revised splitting in place of the original splitting.

During the integrations from day 0 to day 3600, the meridional mass transport was diagnosed at 15-day intervals. Versions (b) and (c) of the model, which use the original splitting, show considerable time variations in the net (vertically summed) transport. The general nature and amplitude of the oscillations are similar in the two models. Similar results are also obtained with the original splitting, but with $\alpha_0 = 0.97297$ cm^3/g. This is the value of α in layer 8, and it was chosen in order to check the effect of the value of α_0 on the behavior of the original splitting. The computation for this case used the same time smoothing as with the revised splitting. In the 15-day samples described here, the oscillations obtained with the revised splitting are generally much smaller than those obtained with the original splitting.

This phenomenon was investigated more closely by integrating the system beyond day 3600 and taking diagnostics of the mass transport at each baroclinic time step. Vertically summed transports were then plotted as follows. For each diagnostic time, compute the maximum and minimum of the transports for that time over all latitudes. Positive transport is taken to be northward, and negative transport is southward. Due to the solid-wall boundary conditions at the northern and southern boundaries, the maximum is at least zero, and the minimum is at most zero. Then plot the maximum as a function of time and the minimum as a function of time. If the model were at a steady state, the maximum and minimum would be identically zero. Deviations from zero indicate a net north–south movement of fluid.

The top frame in Fig. 6 illustrates the net transport obtained with the revised splitting from day 3600 to day 3645. In this plot, the vertical coordinate is the mass transport in Sverdrups (Sv), where 1 Sv $= 10^6$ m^3/s. Over this time interval, the system displays an oscillation with a period of approximately 15 h (24 oscillations in 15 days). As seen below, the oscillation is found mainly in a range of latitudes from approximately 35° N to 40° N. The periods of inertial oscillations at latitudes 35° and 40° are approximately 20.9 and 18.7 h, respectively [4, 13]. The lower frame shows the net transport obtained with the original splitting, as implemented in version (c) of the model. In this case the oscillations are much larger than with the revised splitting, and they have a period of approximately 40 h. As seen later, these oscillations are more distributed with respect to latitude.

The question arises as to whether these oscillations are of physical origin or whether they represent the response of the system to forcing, resulting from numerical discretization errors. First of all, the large discrepancy between the two cases indicates that numerical effects must play a large role in the oscillations in at least one of these cases. Further information is given by an additional experiment in which the barotropic and baroclinic time steps were each cut in half, and the computations were restarted from day 3600. For both the revised and original splittings, the sloshing patterns drifted away from those obtained with the original time steps. Figure 7 shows plots of the meridional transports from day 3625 to day 3645. In the lower frame, the solid curve shows the transport produced with the original splitting and the original time steps, and the dashed figure shows the results

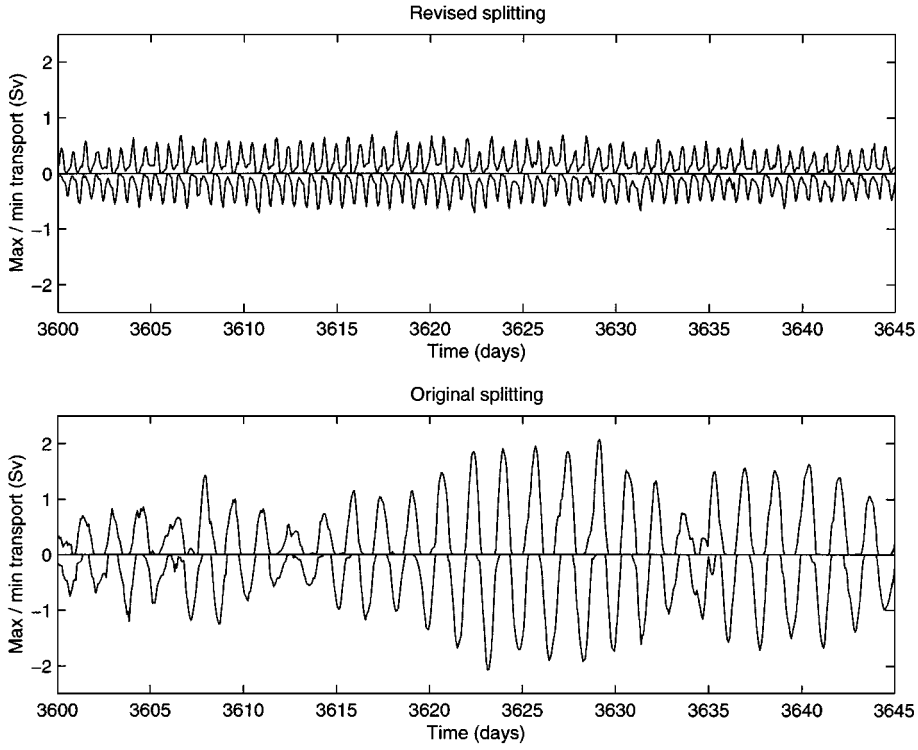

FIG. 6. Illustration of north–south mass transports obtained with versions (a) and (c). In each graph, at each time, the vertical coordinate of the upper curve is the maximum net north–south transport, where the maximum is taken over a discrete set of latitudes. The vertical coordinate of the lower curve is the minimum rate of transport. Positive rates correspond to northward transport, and negative rates correspond to southward transport. Rates are expressed in Sverdrups (Sv), where 1 Sv $= 10^6$ m^3/s. The plots show transports from day 3600 through 3645. Samples are taken at every baroclinic time step, with 16 steps per day.

obtained with the smaller time steps. The upper frame shows analogous results for the revised splitting; in this case both graphs are shown as solid curves for the sake of visibility. In the case of the original splitting, the oscillations appear to be well resolved with respect to time when the longer time step is used, yet reducing the time step produces major changes in the results. This suggests strongly that the oscillations are largely a numerical artifact. Using the revised splitting reduces the oscillations substantially. However, the discrepancy between the two curves plotted in the upper frame suggests that numerical effects may play a significant role in this case as well.

The preceding discussion has been concerned with the net lateral transport obtained by summing over all layers, but the sloshing pattern can also be seen in the depth-dependent circulation. The top frame in Fig. 8 shows a meridional streamfunction obtained with the revised splitting. As before, the mass transports have been summed zonally. In this plot, the vertical coordinate is expressed in terms of fluid layers, instead of z, and the horizontal coordinate is the latitude, with the northernmost latitude of the right. The contour interval is 1 Sv. The plot shows a time average over every baroclinic time step from day 3600 to day 3615, with 16 steps per day. The lower frame in Fig. 8 displays the standard deviation of this time series, plotted as a function of latitude and depth. In the lower frame the contour interval is 0.2 Sv. The plot shows that the sloshing pattern varies with depth and is confined

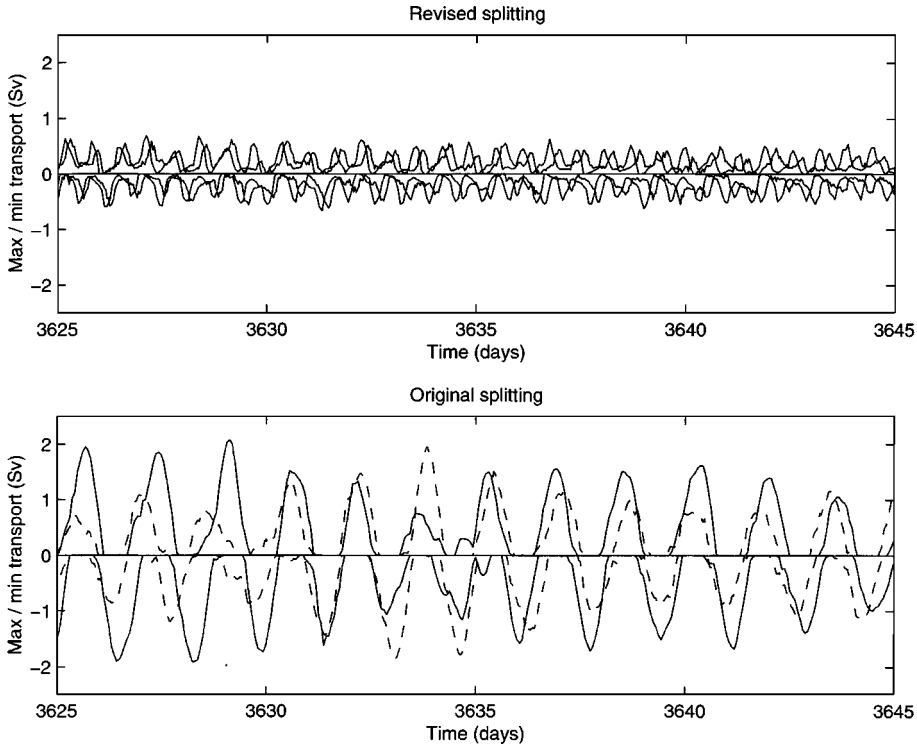

FIG. 7. Effects of the time steps on the north–south mass transport. In the lower frame, the solid curves represent the maximum and minimum transports obtained with the original splitting and with the same time steps used previously. The dashed curves show the transports obtained when the computation is restarted at day 3600 and carried forward with the baroclinic and barotropic time steps equal to half those used for the solid curve. Transports are shown from day 3625 to day 3645. The upper frame shows analogous results obtained with the revised splitting. In the upper frame, both plots are shown as solid curves for the sake of visibility.

mainly to the range of latitudes mentioned earlier, with a maximum standard deviation of roughly 0.8 Sv.

Figure 9 shows analogous results for the original splitting, as implemented in version (c). The lower frame shows a sloshing pattern in approximately the same location as the one obtained with the revised splitting, but the maximum standard deviation is approximately 1.4 Sv in the present case. There is another localized event centered at 20° S, with a maximum standard deviation of approximately 1.0 Sv. Standard deviations in the range 0.2 to 0.4 Sv are also seen throughout much of the model domain.

Additional diagnostics are given by the values of the residual forcing terms $\mathbf{G}$ and $\mathbf{G}^*$ in the vertically averaged barotropic momentum equations (2.3) and (2.5). In these equations, the residual terms represent the combined effects of quantities that are not represented by explicit formulas. In each case, the residual term is taken to be independent of time over each baroclinic time interval, and its magnitude gives a measure of the inexactness in the splitting. In the case of the original splitting, the term $\mathbf{G}^*$ includes the difference between $-\alpha_0 \nabla (p_b' \eta)$ and $-\overline{\nabla M}$. As noted at the end of Section 3.1.2, this difference involves quantities that vary on the fast barotropic time scale. Using the revised splitting has the effect of pulling this difference out of the residual term and including it in the quantities that are represented with explicit formulas.

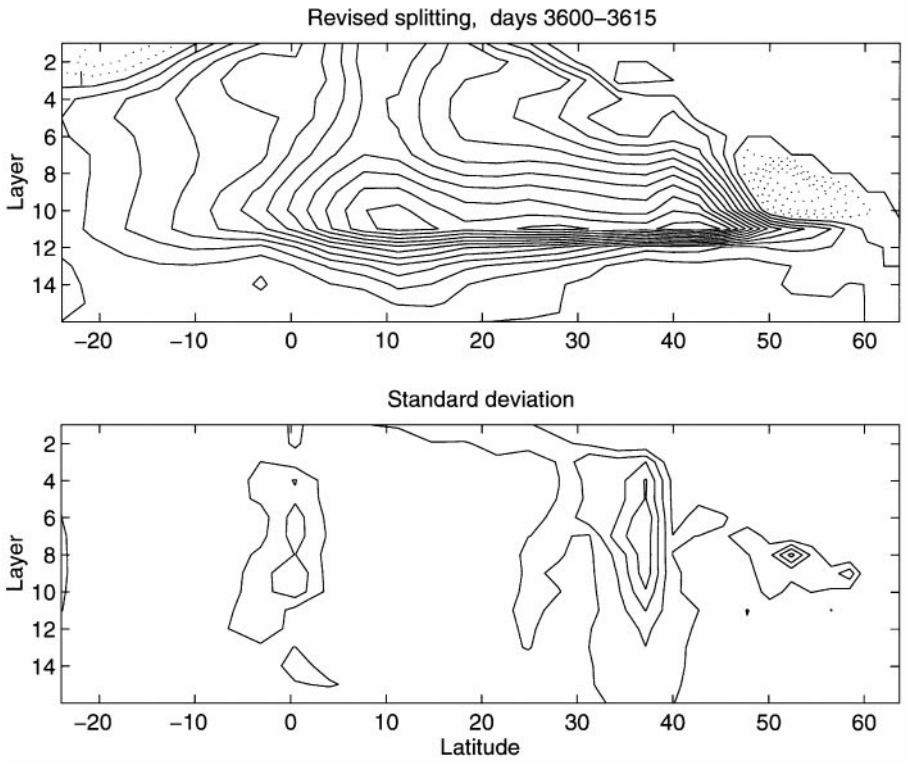

FIG. 8. Depth-dependent circulation obtained with the revised splitting, as implemented in version (a). The upper frame shows a meridional streamfunction obtained from an east–west sum of north–south transports in each layer. The horizontal coordinate is latitude, and the vertical coordinate is the layer. The contour interval is 1 Sv. The plot shows a time average over every baroclinic time step from day 3600 to day 3615, with 16 steps per day. The lower frame shows the standard deviation of this time series, plotted as a function of latitude and depth. In the lower frame the contour interval is 0.2 Sv.

Figure 10 shows values of the components of the residual at day 3600 at latitude 37° N. In each plot the horizontal coordinate is longitude, and the vertical coordinate has units cm/s^2. The upper left and upper right plots show the north–south and east–west components, respectively, of the residual **G** that is produced with the revised splitting. The lower row shows the components of $\mathbf{G}^*$ that are produced by the original splitting, as implemented in version (c). The plots indicate that the revised splitting produces much smaller residuals.

The dependence of the residuals on time is indicated in Fig. 11. The plots in the upper row show values of the north–south component of **G** that are obtained with the revised splitting from day 3600 to day 3615 with 16 baroclinic time steps per day. The time series are taken at positions 37° N 63° W and 37° N 26° W. These positions are located, respectively, at the southern edge of the Gulf Stream and in a relatively quiescent region of the eastern Atlantic. Each plot shows a constant general trend plus perturbations that are seemingly random. In each graph, the solid curve represents data obtained with the model as described previously, and the dotted curve shows data obtained when the code was modified so as to compute **G** in double precision with all other aspects of the model unchanged. The computation of **G** involves the summation of numbers of varying sign which nearly cancel, and the discrepancies between the solid and dotted curves suggest the effects of finite precision. However, finite precision does not give a complete description of the rapid perturbations, as

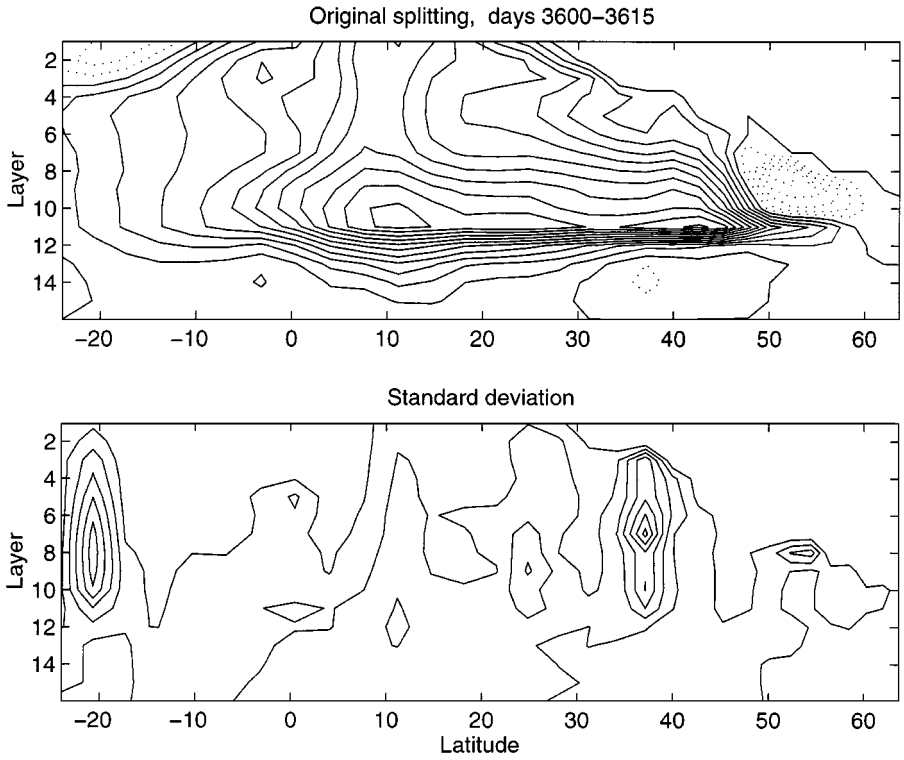

FIG. 9. Depth-dependent circulation obtained with the original splitting, as implemented in version (c). The format is the same as in Fig. 8. The lower frame indicates a greater amount of north–south sloshing than is seen with the revised splitting.

such perturbations are not found in the corresponding plots for the original splitting, which are discussed below. In the upper left frame the average of the values of **G** is approximately 6×10^{-6}, and in the upper right frame the average is approximately 5×10^{-6}. A plot analogous to Fig. 6 (not shown here) shows that the double-precision computation of **G** has very little effect on the maximal and minimal meridional mass transport.

The lower frames in Fig. 11 show the north–south components of $\mathbf{G}^*$ that are obtained with the original splitting at the same locations. These plots show smooth oscillations with a period of approximately one day. However, there is no daily cycle in the forcing functions that drive the model, so the oscillations in $\mathbf{G}^*$ are apparently an artifact of the algorithm. The vertical scale in the lower frames is the same as for the upper frames, but the origin is shifted off-scale because of the greater magnitudes of the quantities plotted in the lower frames. These quantities are almost two orders of magnitude larger than the averages of the values seen in the upper frames.

One could develop an exact formula for the residual term, as it is the mass-weighted vertical average of the terms that are not represented explicitly in the barotropic momentum equation. There is then a corresponding error when the residual is taken to be independent of time on each baroclinic time interval. This error can be regarded as a forcing function that drives the barotropic velocity field $\bar{\mathbf{u}}$ away from its correct values, and this forcing may be a mechanism that generates the sloshing phenomenon discussed above. The quantities **G** (or $\mathbf{G}^*$) and $\bar{\mathbf{u}}$ also appear in the baroclinic momentum equation, so the residual can affect the depth-dependent circulation as well. With the revised splitting, the magnitude of

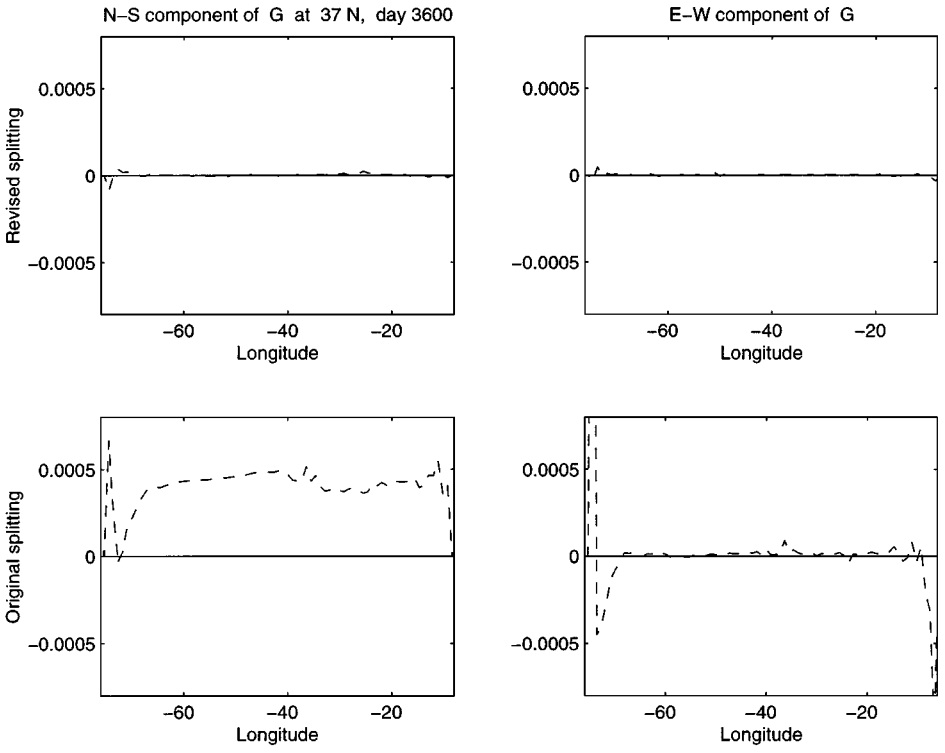

FIG. 10. Components of the residual forcing term in the barotropic momentum equation. The magnitude of this term gives a measure of the inexactness in the splitting. The values shown here are taken at latitude 37° N at day 3600 and are expressed as functions of longitude. The graphs in the upper row of plots show the north–south and east–west components, respectively, of the residual **G** that is obtained with the revised splitting. The graphs in the lower row show the components of $\mathbf{G}^*$ that are obtained with the original splitting, as implemented in version (c).

the residual is greatly reduced, and it appears that this is at least partly responsible for the reduction in the lateral sloshing that is seen with the revised splitting.

5. SUMMARY AND CONCLUSIONS

One of the purposes of this paper is to describe an implementation of the revised splitting that is applicable to an isopycnic ocean circulation model with strongly varying bottom topography. A significant task is to develop a formulation of $\overline{\nabla M}$ that does not include any terms that contain discontinuities that are not already intrinsic to the problem. Such discontinuities can generate spurious gravity waves that interfere with the computed solution and perhaps even cause an ultimate failure in the computation.

A second goal of this work is to test this implementation in the ocean model for which the earlier splitting in [1] was originally developed. In the tests described here, the revised splitting substantially reduces a numerically induced sloshing pattern that is generated by the original splitting. The phenomenon appears related to a residual term that appears in the momentum equations and gives a measure of the accuracy of the splitting.

In linearized model problems, the revised splitting has much better stability properties than the original splitting. This suggests that the revised splitting might make it possible to run the model using less time smoothing than with the original splitting. However, this was

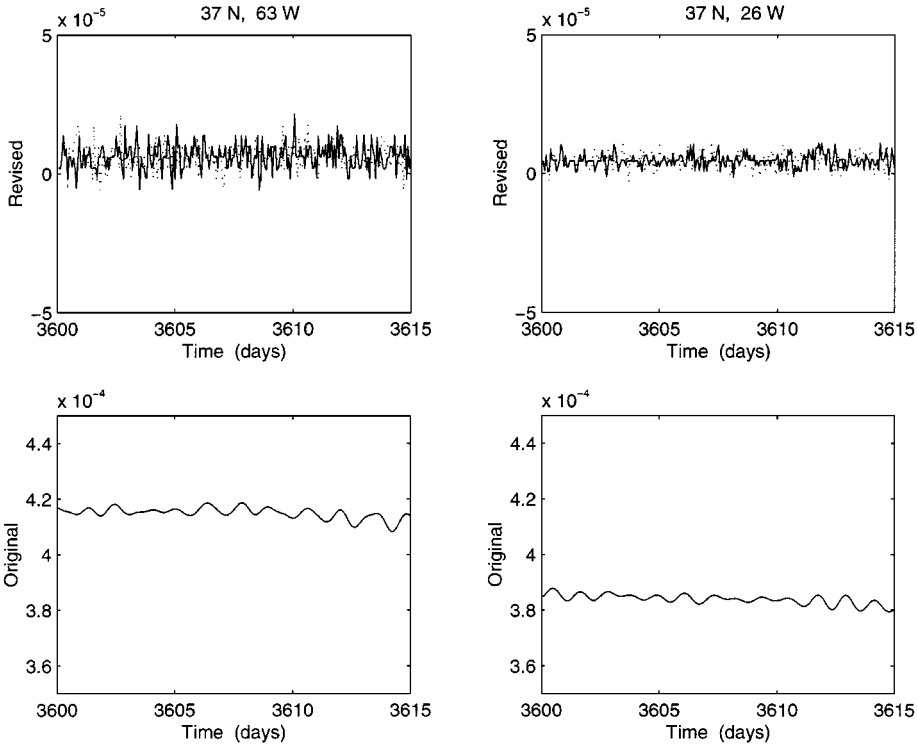

FIG. 11. Time variation of the north–south component of the residual term. The graphs in the upper row show values of **G** obtained with the revised splitting, and the graphs in the lower row show values of $\mathbf{G}^*$ obtained with the original splitting. The vertical ranges in these frames are the same, but the origin is shifted off-scale in the lower two frames due to the greater magnitudes of the quantities plotted there. In the upper row of plots, the solid curves represent values that are obtained when **G** is computed in single precision, and the dotted curves illustrate the results when **G** is computed in double precision with all other aspects of the model unchanged.

not the case in these particular tests. One factor may be that, at least in the linearized setting, the original splitting contains substantial dissipation in some modes, even while there is instability in other modes. Another factor is that the model tested here uses a numerical method for the baroclinic equations which admits a sawtooth computational mode that can be particularly stimulated by the implementation of the mixed layer in that model. This mode must be suppressed, regardless of how the barotropic–baroclinic splitting is done. These remarks suggest the further development of numerical methods, although such investigations are beyond the scope of the present paper.

The computations described here use a grid that does not give high resolution of boundary currents or of eddies that are shed from ocean currents. An examination of the effects of the revised splitting on the modeling of these phenomena is another issue for further investigation.

ACKNOWLEDGMENTS

I thank Rainer Bleck, Sumner Dean, Roland de Szoeke, Scott Springer, and Shannon Wynne for useful conversations on matters related to the subject of this paper. This work was supported by National Science Foundation Grant DMS-9407509 and by the program on Computer Hardware, Advanced Mathematics and Model Physics

of the U.S. Department of Energy. The computations described here used the facilities of the Environmental Computing Center at Oregon State University and the Advanced Computing Laboratory at Los Alamos National Laboratory.

REFERENCES

1. R. Bleck and L. T. Smith, A wind-driven isopycnic coordinate model of the north and equatorial Atlantic Ocean 1. Model development and supporting experiments, *J. Geophys. Res. C* **95**, 3273 (1990).
2. R. Bleck, C. Rooth, D. Hu, and L. T. Smith, Salinity-driven thermocline transients in a wind- and thermohaline-forced isopycnic coordinate model of the North Atlantic, *J. Phys. Oceanogr.* **22**, 1486 (1992).
3. J. K. Dukowicz and R. D. Smith, Implicit free-surface method for the Bryan–Cox–Semtner ocean model, *J. Geophys. Res. C* **99**, 7991 (1994).
4. A. E. Gill, *Atmosphere-Ocean Dynamics* (Academic Press, San Diego, 1982).
5. R. Hallberg, Stable split time stepping schemes for large-scale ocean modeling, *J. Comput. Phys.* **135**, 54 (1997).
6. G. J. Haltiner and R. T. Williams, *Numerical Prediction and Dynamic Meteorology* (Wiley, New York, 1980).
7. R. L. Higdon and A. F. Bennett, Stability analysis of operator splitting for large-scale ocean modeling, *J. Comput. Phys.* **123**, 311 (1996).
8. R. L. Higdon and R. A. de Szoeke, Barotropic-baroclinic time splitting for ocean circulation modeling, *J. Comput. Phys.* **135**, 30 (1997).
9. Y.-J. Hsu and A. Arakawa, Numerical modeling of the atmosphere with an isentropic vertical coordinate, *Monthly Weather Rev.* **118**, 1933 (1990).
10. P. D. Killworth, D. Stainforth, D. J. Webb, and S. M. Paterson, The development of a free-surface Bryan–Cox–Semtner ocean model, *J. Phys. Oceanogr.* **21**, 1333 (1991).
11. F. Mesinger and A. Arakawa, *Numerical Methods Used in Atmospheric Models*, GARP Publ. Ser. No. 17, Vol. 1 (WMO-ICSU Joint Organizing Committee, Geneva, 1976).
12. J. M. Oberhuber, Simulation of the Atlantic circulation with a coupled sea ice—mixed layer—isopycnal general circulation model. Part I. Model description, *J. Phys. Oceanogr.* **23**, 808 (1993).
13. J. Pedlosky, *Geophysical Fluid Dynamics*, 2nd ed. (Springer-Verlag, New York, 1987).
14. A. J. Semtner, Finite-difference formulation of a world ocean model, in *Advanced Physical Oceanographic Numerical Modelling*, edited by J. J. O'Brien (Reidel, Norwell, MA, 1986), p. 187.