High-Order Monotonic Numerical Diffusion and Smoothing

MING XUE

Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma

(Manuscript received 28 May 1999, in final form 17 September 1999)

ABSTRACT

High-order numerical diffusion is commonly used in numerical models to provide scale selective control over small-scale noise. Conventional high-order schemes have undesirable side effects, however: they can introduce noise themselves. Two types of monotonic high-order diffusion schemes are proposed. One is based on flux correction/limiting on the corrective fluxes, which is the difference between a high-order (fourth order and above) diffusion scheme and a lower-order (typically second order) one. Overshooting and undershooting found in the solutions of higher-order diffusions near sharp gradients are prevented, while the highly selective property of damping is retained.

The second simpler (flux limited) scheme simply ensures that the diffusive fluxes are always downgradient; otherwise, the fluxes are set to zero. This much simpler scheme yields as good a solution in 1D cases as and better solutions in 2D than the one using the first more elaborate flux limiter. The scheme also preserves monotonicity in the solutions and is computational much more efficient.

The simple flux-limited fourth- and sixth-order diffusion schemes are also applied to thermal bubble convection. It is shown that overshooting and undershooting are consistently smaller when the flux-limited version of the high-order diffusion is used, no matter whether the advection scheme is monotonic or not. This conclusion applies to both scalar and momentum fields. Higher-order monotonic diffusion works better and even more so when used together with monotonic advection.

1. Introduction

Most numerical models employ numerical diffusion or computational mixing to control small-scale (near two grid intervals in wavelength) noise that can arise from numerical dispersion, nonlinear instability, discontinuous physical processes, and external forcing. Such diffusion cannot always be substituted for by physical subgrid-scale turbulence parameterization, whose coefficients can go to zero in stable flows where numerical noise can still arise. Discussions on the use of diffusion or filter to remove small-scale noise can be found in, for example, Shapiro (1970, 1975), Raymond and Garder (1976), and Raymond (1988).

Diffusion is often introduced by adding an additional term to the right-hand side of the time-dependent equations, and it takes on different forms, as shown in Eq. (1):

$$\frac{\partial \phi}{\partial t} = S + (-1)^{n/2+1}\alpha_n \nabla^n \phi, \tag{1}$$

where ϕ is any prognostic variable. The S in (1) represents other processes or sources. The last term on the right-hand side (rhs) is the added diffusion term and n ($=0, 2, 4, 6, \ldots$) denotes the order. Here α_n is called diffusion coefficient. Diffusion or smoothing can also be introduced by periodically applying a spatial filter or smoother to the predicted field (e.g., Shapiro 1970, 1975). Its effect is often equivalent to applying diffusion in the prognostic equation for a period of time.

Second- and fourth-order formulations are most commonly used in gridpoint models (e.g., Klemp and Wilhelmson 1978; Xue et al. 1995). Formulations of higher order are often used in spectral models where their implementation is as simple as for the lower-order ones (see discussion by Sardeshmukh and Hoskins 1984; Hoskins 1980). Zero-order ($n = 0$) damping is also used in models, but usually only in the lateral and/or upper boundary zones when nonselective relaxation is needed. The case for using high-order scale-selective filters has also been made by Shapiro (1975) and Raymond (1988).

The second-order diffusion term ($n = 2$) has a Laplacian form and resembles the physical diffusion/heat transfer term in which higher values are always diffused/transferred to regions of smaller values. As a result, it never creates new extrema not present previously in the field of the prognostic variable. The second-order as well as zero-order diffusion is not as selective as higher-order formulations, however, in that waves longer than

Corresponding author address: Dr. Ming Xue, Center for Analysis and Prediction of Storms, University of Oklahoma, 100 E. Boyd, Norman, OK 73019.
E-mail: mxue@ou.edu

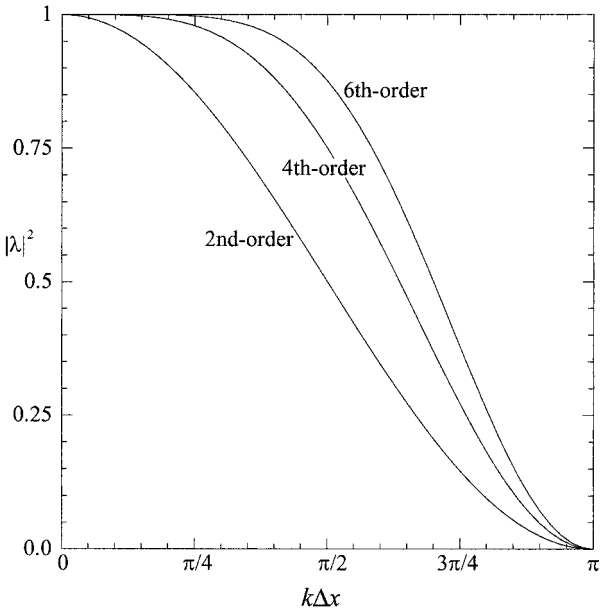

FIG. 1. Amplification factors $|\lambda|^2$ as given in Eq. (2) plotted against wavenumber ($k\Delta x$) for second-, fourth-, and sixth-order diffusion schemes. For each scheme, $\alpha\Delta t$ is chosen so that $|\lambda| = 0$ for $k\Delta x = \pi$, which corresponds to a wavelength of $2\Delta x$, i.e., the shortest waves representable on a discrete grid. In another word, $2\Delta x$ waves a completely damped after one time step of integration.

two grid intervals tend to be damped significantly. Higher-order diffusion damps waves more selectively.

When the one-dimensional version of Eq. (1) is solved using a forward in time and centered in space finite difference scheme, the linear stability analysis using the von Neumann method (e.g., Haltiner and Williams 1980) yields the following amplification (damping in this case) factors for the four lowest (zero, second, fourth and sixth) order schemes:

$$|\lambda|^2 = 1 - \alpha\Delta t\begin{cases}1\\ [2 - 2\cos(k\Delta x)]/\Delta x^2\\ [6 - 8\cos(k\Delta x) + 2\cos(2k\Delta x)]/\Delta x^4\\ [20 - 30\cos(k\Delta x) + 12\cos(2k\Delta x)\\ \quad - 2\cos(3k\Delta x)]/\Delta x^6],\end{cases} \tag{2}$$

where Δt is the integration time step size, Δx the horizontal grid space, and k the wavenumber. For the time integration to be stable, $|\lambda|$ has to be no greater than 1, and the condition can be met by choosing sufficiently small Δt or α. Figure 1 shows $|\lambda|^2$ as a function of $k\Delta x$ with $\alpha\Delta t$ chosen for each scheme so that $|\lambda| = 0$ for $k\Delta x = \pi$, which corresponds to a wavelength of $2\Delta x$, the shortest waves that can be represented on a discrete grid. In another word, $2\Delta x$ waves are completely damped in one time step. (Note that in practice one rarely chooses as large a diffusion coefficient. When one does, the effect of diffusion in each time step is equivalent to one application of a Shapiro-type spatial filter, which is rarely applied every time step, however. When a smaller diffusion coefficient is chosen, significant reduction in the amplitude of $2\Delta x$ waves occurs in a number of time steps with the diffusion effect being accumulative).

It can be seen from Fig. 1 that for a four-grid interval wave, the amplitude is reduced by half by a second-order scheme but only 20% by sixth-order scheme, while the amplitude of two grid interval waves ($k\Delta x = \pi$) is reduced to zero in all cases. Clearly higher-order schemes are more scale selective.

Higher-order schemes have undesirable side effects, however. They can introduce spatial oscillations known as the Gibbs phenomenon, as evidenced by Fig. 2. (Note that the Gibbs phenomenon can also be introduced by most conventional advection schemes. We shall limit our attention here to the diffusion process, however.) The figure shows the numerical solutions of Eq. (1) (without S) for $n = 2$, 4 and 6, i.e., for second-, fourth-, and sixth-order diffusion, respectively. Forward in time and centered in space finite difference scheme is used for the time integration. The 1D x-domain width is 1 represented by 50 grid zones, that is, $\Delta x = 1/50$. Further, $\alpha_n = (\Delta t)^{-1}(\Delta x/2)^n$ and $\Delta t = 1$ are chosen so that $|\lambda| = 0$ for $k\Delta x = \pi$, that is, for $2\Delta x$ waves. Therefore, two-grid interval waves are completely damped within one time step for all schemes. In the initial profile (thick line), ϕ jumps from -0.4 to 0.4 within one grid interval. Consecutive solutions at time intervals of 10 are plotted as thin lines, with the last one being at $t = 100$. Periodic lateral boundary conditions are used.

It can be seen that the second-order scheme is strongly diffusive even for broad-scale features, and amplitude of the entire square wave is significantly reduced at the end of time integration ($t = 100$). The fourth- and sixth-order schemes only modify the solution near the sharp gradient, where short waves dominate. A sharper gradient is maintained in the solution of the latter, while both keep the solution at the center of the square wave unchanged. The over- and undershoot in the solutions are evident, and both positive and negative oscillations are present in the fourth- and sixth-order solutions. They are clearly undesirable.

Sardeshmukh and Hoskins (1984) showed that when Eq. (1) is solved in a spectral space, the effect of the diffusion term after a given period of time is equivalent to applying a spatial filter, whose result at each point is a weighted average of itself and values at the surrounding points. The weighting function is not all positive for high-order schemes, which explains the source of the Gibbs oscillations. An additional physical explanation is that with high-order schemes, the diffusive flux is no longer proportional to local gradient and does not necessarily point in the downgradient direction. There is no such problem with the second-order diffusion, however. It is obviously desirable to design a class of diffusion schemes or spatial filters that possess both the scale-selective properties of the higher schemes and the

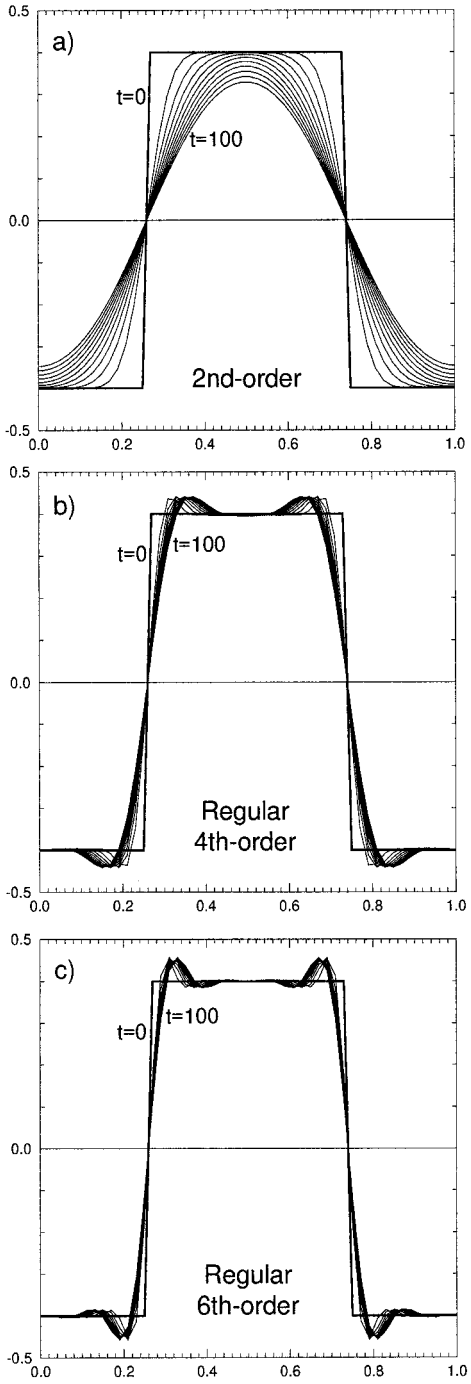

FIG. 2. Numerical solutions of Eq. (1) (without S) for (a) $n = 2$, (b) $n = 4$ and (c) $n = 6$, i.e., for second-, fourth-, and sixth-order diffusion, respectively. The 1D x-domain width is 1 represented by 50 grid zones, i.e., $\Delta x = 1/50$. $\alpha_n = (\Delta t)^{-1}(\Delta x/2)^n$ and $\Delta t = 1$ are chosen so that $|\lambda| = 0$ for $k\Delta x = \pi$, i.e., for $2\Delta x$ waves. Therefore, two grid interval waves are completely damped within one time step for all schemes. In the initial profile (thick line), ϕ jumps from -0.4 to 0.4 within one grid interval. Consecutive solutions at time interval of 10 are plotted as thin lines, with the last one being at $t = 100$. Forward in time and centered in space finite difference scheme is used for the time integration. Periodic lateral boundary conditions are used.

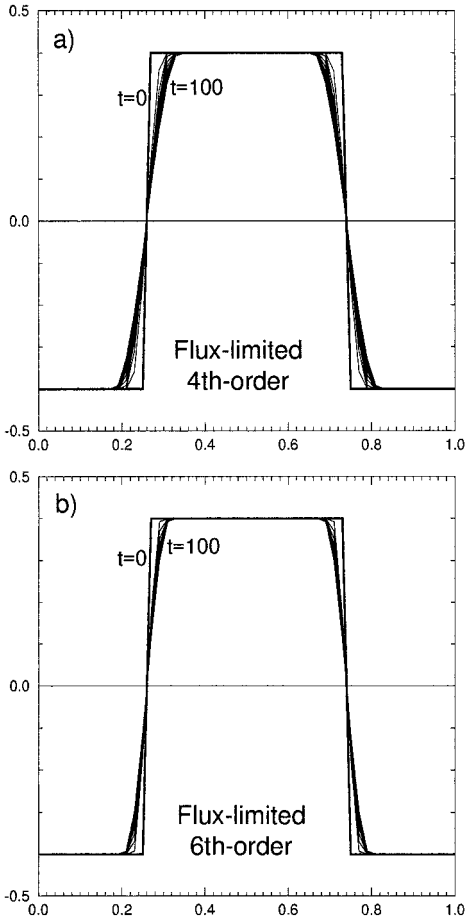

FIG. 3. As in Fig. 2, but for solutions of Zalesak-type flux-limited (a) fourth-order and (b) sixth-order diffusion schemes, with the lower-order scheme used being the second order. Compare (a) to Fig. 2b and (b) to Fig. 2c.

monotonic nature of lower-order schemes. By being monotonic, we mean that the scheme will not create new extrema that are not present previously in the field, nor will it intensify existing extrema. Therefore the solution will be free of Gibbs oscillations. In this paper, we present two types of monotonic schemes, one based on the combination of high- and lower-order schemes with correction to the difference between their fluxes, and the other based on ensuring that the diffusive fluxes are downgradient.

2. Monotonic diffusion with Zalesak-type flux correction

For simplicity, we will present our case in one dimension and note that the method can be readily extended to multidimensions. A one-dimensional flux limiter that ensures monotonicity by advection was first proposed by Boris and Book in the early 1970s (Boris and Book 1973). Zalesak (1979) later proposed a flux-limiting scheme that is truly multidimensional. It is the

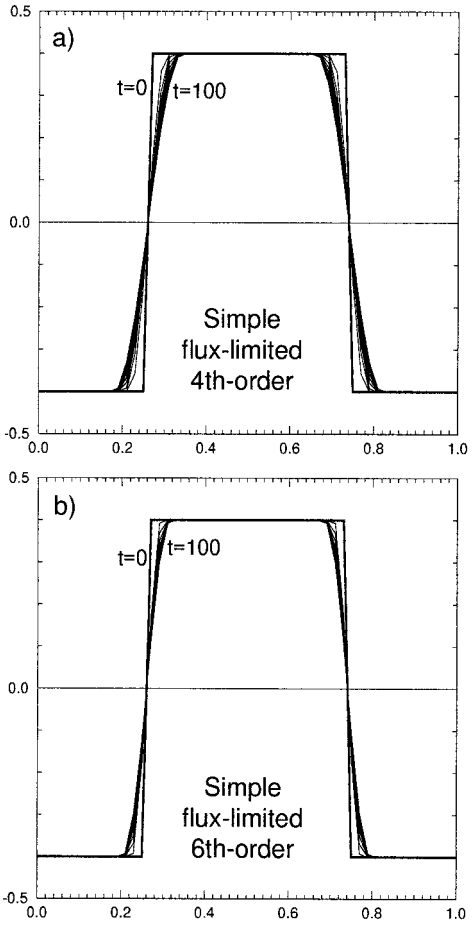

FIG. 4. As in Fig. 3, but for solutions of the simple flux-limited (a) fourth-order and (b) sixth-order diffusion schemes, in which upgradient diffusive fluxes are set to zero. Compare (a) to Fig. 2a and (b) to Fig. 2b.

latter on which we base our first monotonic higher-order diffusion scheme. The limiter of Zalesak is also used in the monotonic option of the multidimensional positive definite advection transport scheme of Smolarkiewicz and Grabowski (1990).

Without the source term, the 1D version of Eq. (1) can be written as

$$\frac{\partial \phi}{\partial t} = -\frac{\partial}{\partial x}\left[(-1)^{n/2}\alpha_n \frac{\partial^{n-1}\phi}{\partial x^{n-1}}\right] \equiv -\frac{\partial F}{\partial x}. \tag{3}$$

We call $-F$ the diffusive flux. Using notations

$$\delta_x \psi = (\psi_{i+1/2} - \psi_{i-1/2})/\Delta x,$$

$$\delta_{xx}\psi = \delta_x(\delta_x \psi) = (\psi_{i+1} - 2\psi_i + \psi_{i-1})/\Delta x^2$$

and so on, we integrate Eq. (3) using a forward in time and centered in space scheme,

$$\phi^{n+1} = \phi^n - (A^n_{i+1/2} - A^n_{i-1/2}), \tag{4}$$

where n denotes time level and

$$A^n_{i+1/2} \equiv F^n_{i+1/2}\Delta t/\Delta x.$$

These are time integration schemes for which the amplification factors in Eq. (2) are derived, and they are all conditionally stable. In the above, we assume the grid spacing Δx is constant.

Following the flux correction idea of Zalesak (1979) that was applied to advection, we propose here a flux corrected scheme based on the combination of lower-order and a high-order diffusion schemes. We rewrite Eq. (4) as

$$\phi_i^* = \phi_i^n - (A^L_{i+1/2} - A^L_{i-1/2}),$$

$$\phi_i^{n+1} = \phi_i^* - (C_{i+1/2}A^{\mathrm{HL}}_{i+1/2} - C_{i-1/2}A^{\mathrm{HL}}_{i-1/2}), \tag{5}$$

where A^L is a lower-order diffusive flux that we choose to be of second order (because it is the highest-order simple diffusion scheme that is still monotonic). Here A^{HL} is the difference between a high-order diffusive flux A^H and the low-order flux A^L, that is,

$$A^{\mathrm{HL}} = A^H - A^L.$$

We call A^{HL} the corrective flux that attempts to improve the lower-order solution ϕ^*.

The factor C is the flux correction coefficient to be determined for each grid point; it is also called the flux limiter. It is clear that when $C = 0$, Eq. (5) reduces to the lower-order scheme (which is not preferred because of its diffusive effect on longer waves), while when $C = 1$, it becomes purely high order (which is preferable due to its scale selectivity but can produce ripples at places). Typically $0 \le C \le 1$. The idea is to use the high-order scheme to the maximum extent possible (so that waves longer than two grid intervals are less affected) while satisfying certain monotonicity constraints.

We should point out here when the lower-order diffusive flux is used together with the high-order one, their coefficients, α_n, should be chosen to be consistent with each other. To do so, we require that two grid interval waves be damped by the same amount for a given number of time steps (longer waves will be damped by a different amount, however). It should also be noted that unlike the Zalesak flux-corrected transport (FCT) scheme for advection where both lower- and high-order schemes are numerically consistent approximations to the same advection term, our lower- and high-order schemes are two different formulations both aimed at damping unphysical noise.

Since a monotonic scheme should not create new extrema in the solution, we require

$$\phi_i^{\min} \le \phi_i^{n+1} \le \phi_i^{\max}. \tag{6}$$

Since we know the diffusion process does not generate new extrema, we choose $\phi_i^{\min}$ and $\phi_i^{\max}$ as the minimum and maximum at the current and the immediate neighboring points at the previous time step, that is,

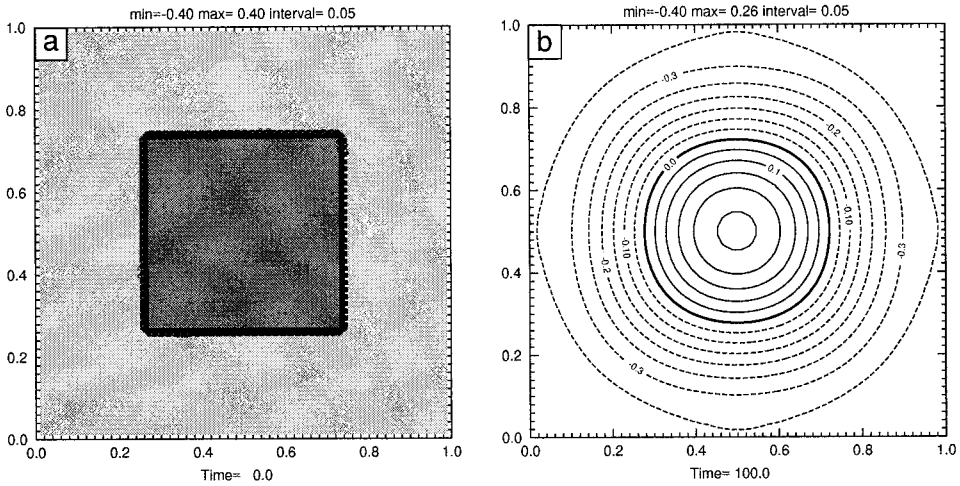

FIG. 5. Contour maps of a 2D field at (a) initial time, and (b) after subjecting to second-order diffusion for 100 time steps at time 100. Negative contours are dashed. Dark shading represent values equal to or greater than 0.4, while light shading represents values equal to or smaller than −0.4. No such region exists in (b) because the field is significant diffused. The shading indicates areas of possible overshooting or undershooting. The same parameters used by 1D experiments (cf. Fig. 2) are used here, with the only difference being the second dimension.

$$\phi_i^{\min} = \min(\phi_{i-1}^n, \phi_i^n, \phi_{i+1}^n),$$
$$\phi_i^{\max} = \max(\phi_{i-1}^n, \phi_i^n, \phi_{i+1}^n). \quad (7)$$

For advective process, Zalesak (1979) also include the minimum and maximum in the current-level solution by the lower-order advection. Since the second-order diffusion always reduces the amplitude, the inclusion of the lower-order solution is not necessary here.

With $\phi_i^{\min}$ and $\phi_i^{\max}$ properly defined, we now need to find the flux correction coefficient C that will ensure that condition (6) is satisfied. Following Zalesak (1979), we choose C to be

$$C_{i+1/2} = \begin{cases} \min(1, R_{i+1}^+, R_i^-) & \text{if } A_{i+1/2}^{\mathrm{HL}} \geq 0 \\ \min(1, R_i^+, R_{i+1}^-) & \text{if } A_{i+1/2}^{\mathrm{HL}} < 0, \end{cases} \quad (8)$$

where

$$R_i^+ = (\phi_i^{\max} - \phi_i^*)/(A_i^{\mathrm{IN}} + \varepsilon),$$
$$R_i^- = (\phi_i^{\min} - \phi_i^*)/(A_i^{\mathrm{OUT}} + \varepsilon). \quad (9)$$

Here A_i^{IN} and A_i^{OUT} are, respectively, the sum of all corrective fluxes (A^{HL}) going *into* and *away from* grid point i, and both are always positive. Here the corrective fluxes can be from one, two, or three directions, depending on the dimensionality of the problem. The ε is a small positive value typically chosen to be 10^{-20} to avoid division by zero in the computer program. It can be shown that by applying limiting factor C to the corrective fluxes, the new solution ϕ^{n+1} obtained in (5) is prevented from falling below $\phi^{\min}$ or overshooting above $\phi^{\max}$. Further explanation of such a limiter can be found in Zalesak (1979) or in Smolarkiewicz and Grabowski (1990).

3. Alternative flux-limited high-order schemes

The flux limiter described in the previous section was designed based on the monotonicity consideration. It was also pointed out in the introduction that diffusive fluxes in high-order schemes are not necessarily, as the second-order one is, downgradient. This can result in unphysical solutions. What if we find a way of preventing upgradient diffusion from happening? A simple way is to set the diffusive flux [$-F$ in Eq. (3)] to zero whenever it has a different sign than the gradient of ϕ; that is, we set

$$-F_{i+1/2} = -F_{i+1/2} \max[0, \mathrm{sign}(-F_{i+1/2}\delta_x\phi)], \quad (10)$$

where the sign( ) operator returns the sign of the operant. Since Eq. (10) also limits the original diffusive flux according to certain criterion, we call Eq. (10) a flux limiter too, albeit a simple one. The solutions of the diffusion equation with this flux-limiter applied to the fourth- and sixth-order schemes are shown in Fig. 4. Interestingly, the results in Fig. 4a are as good as those shown in Fig. 3a, and those in Fig. 4b as good as in Fig. 3b, suggesting a simple flux limiter like that in Eq. (10) may be all one needs to ensure monotonicity of high-order diffusion schemes. This simple limiter is far more efficient than the previous one since the extra computation in Eq. (10) is almost trivial. In the later text, we shall refer to the current class of schemes as the *simple flux-limited* or *simple monotonic* diffusion schemes. The efficacy of these and those based on Zalasak flux limiter in two dimensions will be examined in the next section.

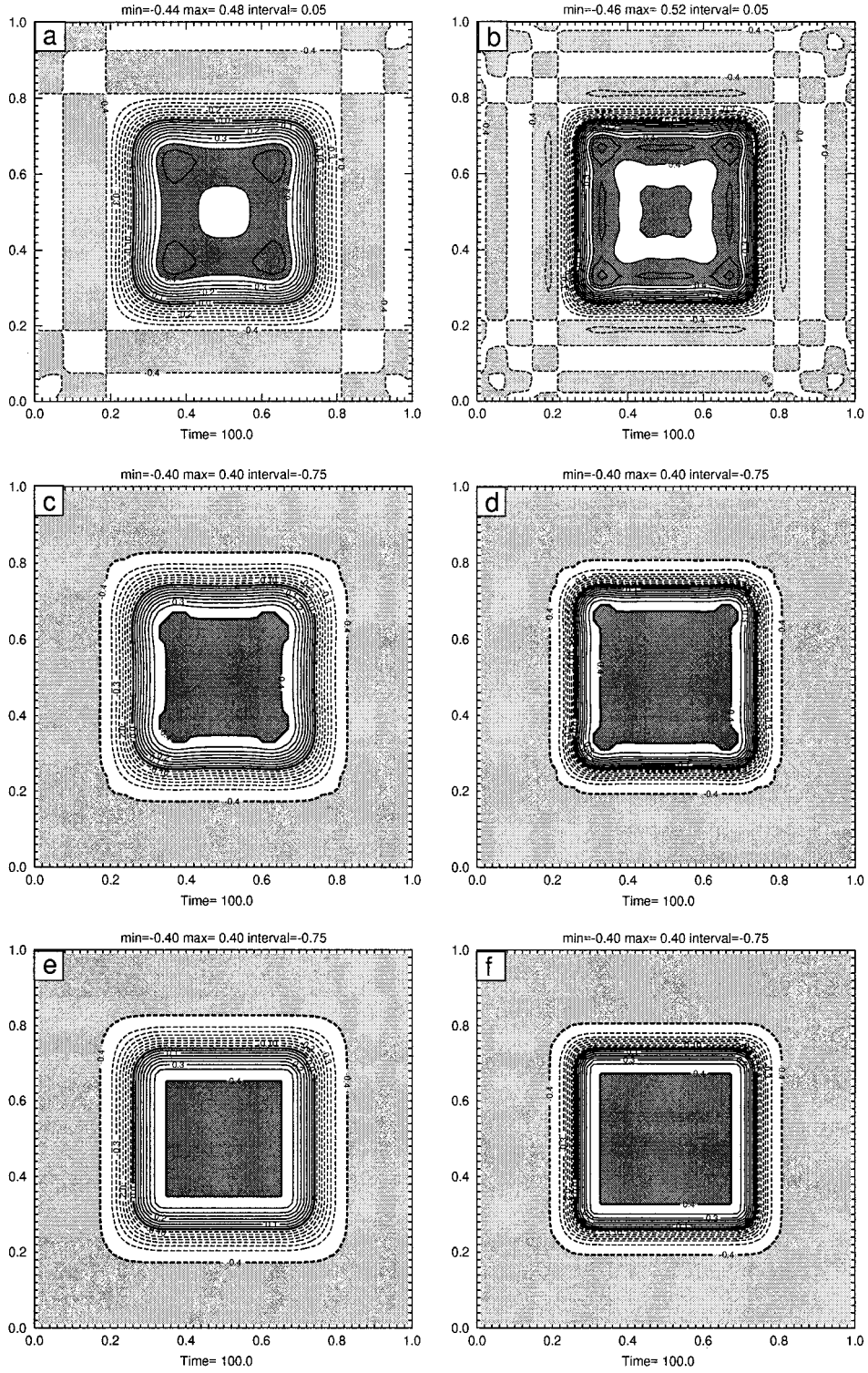

FIG. 6. Contour maps of a 2D field after subjecting to (a) regular fourth-order, (b) regular sixth-order, (c) fourth-order flux-corrected, (d) sixth-order flux-corrected, (e) fourth-order with simple flux limiter, and (f) sixth-order with simple flux limiter diffusion schemes for 100 time steps at time 100. The plotting conventions are the same as in Fig. 5.

TABLE 1. List of bubble experiments together with rms and maximum absolute errors compared with the reference run.

Experiment	Configuration			Rmse	Max abs err
	θ advection	u, w diffusion	θ diffusion		
REF	FCT	Mono sixth order	None	0	0
A4D4	LT4CS	Reg. fourth order	Reg. fourth order	0.1072	0.5270
A4MD4	LT4CS	Mono fourth order	Mono sixth order	0.0926	0.5012
A4D6	LT4CS	Reg. sixth order	Reg. sixth order	0.0781	0.5744
A4MD6	LT4CS	Mono sixth order	Mono sixth order	0.0609	0.5741
A4dD6	LT4CS	Double reg. sixth order	Double reg. sixth order	0.0815	0.5685
A4dMD6	LT4CS	Double mono sixth order	Double mono sixth order	0.0718	0.5318
FCTD6	FCT	Mono sixth order	Reg. sixth order	0.0740	0.5481
FCTMD6	FCT	Mono sixth order	Mono sixth order	0.0578	0.4944
REF1	FCT	Reg. sixth order	None	0.0435	0.4926

4. Application to two-dimensional square wave

In the previous section, examples have been shown for one-dimensional cases for the two flux-limited diffusion schemes. The extension of both methods to multidimensions is straightforward. We examine their performance for the 2D square wave case here.

Figure 5a shows the contours of the 2D version of the square wave in Fig. 2 at initial time. Inside the black box are constant values of 0.4 and outside constant values of -0.4. Figure 5b shows the same field after being diffused by second-order diffusion for 100 times (as in the 1D case). The same grid spacing, time step size, and mixing coefficients are used for 2D tests as in the 1D square wave ones. In Fig. 5b, the field has been significantly smoothed even at relatively large scales, and the peak value is reduced to 0.26 from 0.4. The results using various high-order schemes are shown in Fig. 6.

Figures 6a and 6b show, respectively, the fields at $t = 100$ using regular fourth- and sixth-order diffusion. While a much sharper gradient is retained, overshooting (values larger than 0.4, dark shading) and undershooting (values less than -0.4, light shading) are obvious, with those of the sixth-order scheme being more serious. There is a tendency for the contours to extend outward at the four corners.

In Figs. 6c and 6d, Zalasak-type flux correction is applied to the fourth- and sixth-order-based schemes, respectively. Undershooting and overshooting are completely avoided with the minimum being exactly -0.4 and the maximum exactly 0.4, just as in the 1D case. A better gradient is retained than with the regular diffusion. However, the extrusion of contours at four corners is not cured by the flux correction. Apparently, due to the effect of the second dimension, the effect of flux limiting at the corner is different from that away from the corner, where the flux limiter is essentially one-dimensional.

When the simple flux limiter is used, the results are, interestingly, better (Figs. 6e and 6f). The squared shape of the contours is retained almost perfectly. While little difference is found in the 1D tests, the 2D results suggest that the performance of the simple flux limiter is superior to the more sophisticated and also much more expensive flux limiter given in section 2a. This means a large improvement can be obtained at little extra cost with high-order diffusion schemes with a simple flux limiter.

5. Application to 2D thermal bubble convection

In the previous sections, two classes of monotonic diffusion schemes were tested with stationary 1D and 2D square waves. With them no effect of advection is taken into account. These experiments do show that the simple flux-limited diffusion schemes are superior to the more sophisticated version in both performance and efficiency. Therefore we choose to examine only the simple flux-limited diffusion schemes in a fully nonlinear time-dependent model.

We examine the performance of these schemes in 2D thermal bubble convection. The experiment setup follows Robert (1993). A circular bubble with a 500-m radius and 0.5 K potential temperature (θ) surplus is initially centered at $x = 250$ m, $z = 260$ m inside a 1 km $\times$ 1 km box. Uniform environmental potential temperature is 303.15 K. The nonhydrostatic compressible Advanced Regional Prediction System (ARPS) model (Xue et al. 1995) in its dry 2D mode is used for the experiment. The ARPS employs the mode-splitting time integration technique of Klemp and Wilhelmson (1978) to deal with fast acoustic waves. The model employs leapfrog time integration scheme for the slow modes and fourth-order centered difference for advection. For scalars including θ there is an option for using the FCT scheme for advection. Since the problem is symmetric, we run the model and show the results in the right half of the problem domain only.

First, we present a benchmark (experiment REF in Table 1) solution obtained using the monotonic FCT advection scheme for θ. The FCT scheme is based on fourth order in space, trapezoidal in time, and upstream-forward schemes following Zalesak (1979). The resultant time integration is essentially forward in time. Because the FCT advection is monotonic, it does not introduce any overshoot and undershoot in θ field; furthermore, no other process exists in the model that

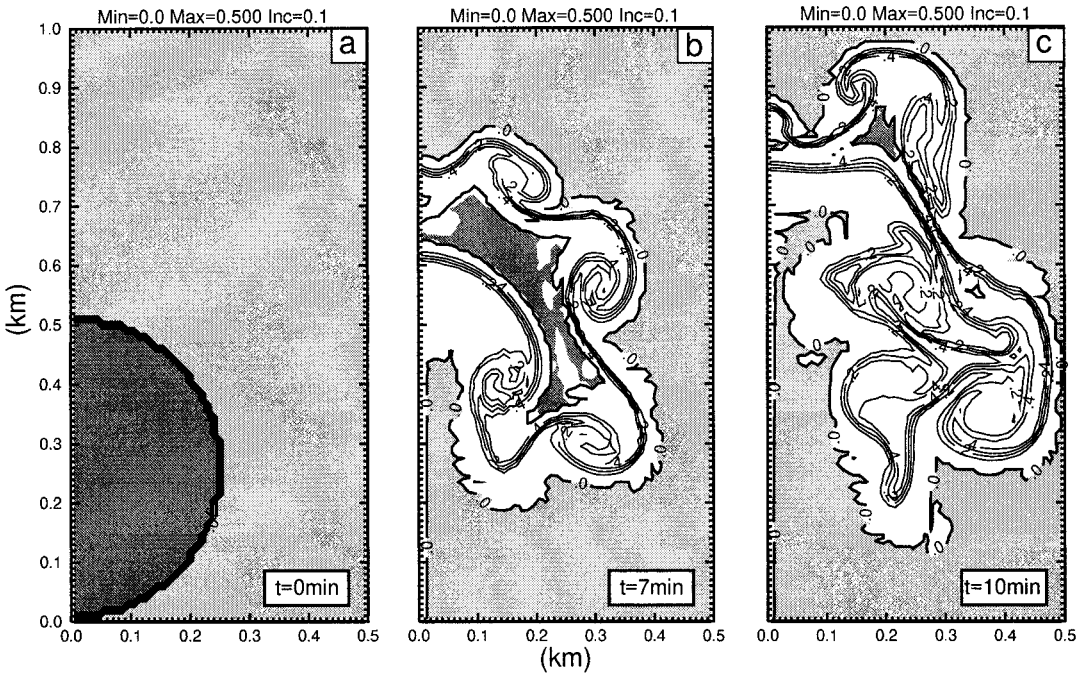

FIG. 7. Perturbation potential temperature (θ') fields at (a) $t = 0$ min, (b) $t = 7$ min, and (c) $t = 10$ min, showing the evolution of a thermal bubble in a neutral environment. Flux-corrected transport is used for potential temperature advection and the simple flux-limited (monotonic) sixth-order diffusion is applied to momentum. Zero and 0.5 contours are thickened, and areas with values equal to or greater than 0.5 are indicated by dark shading whereas areas with values equal to or less than 0.0 are represented by light shading. Hardly any undershooting or undershooting exists in the solution as indicated by the maximum and minimum values in the plots, due to the use of monotonic FCT advection and the absence of nonmonotonic diffusion.

would introduce small-scale oscillations in θ, therefore no explicit diffusion is applied to θ in this benchmark run. A regular leapfrog in time fourth-order centered in space (LT4CS) scheme is used for momentum advection and for momentum a sixth-order *simple* monotonic diffusion is applied. The diffusion on momentum is necessary to control small-scale noise introduced by the nonmonotonic advection [it is difficult to apply a scheme other than the leapfrog centered one to the current compressible system when it is solved using the split-explicit procedure, as discussed by Skamarock and Klemp (1992)]. For numerical stability, the diffusion terms are calculated at the past time level when used together with the leapfrog scheme, but at the present time level when used together with the effectively forward in time FCT scheme. In the former case, superscript n in all derivations in section 2 should be replaced by $n - 1$, and time step size Δt by $2\Delta t$ *if* the Zalesak-type monotonic diffusion scheme is used.

The diffusion coefficient is chosen for this benchmark and all subsequent cases (except for two that used doubled coefficient) to be ¼ of the value by which two grid interval waves are completely damped in one time step; therefore two grid interval waves are damped in approximately four time steps. The grid spacing is 10 m in both directions and the time step size is 0.5 s.

The perturbation potential temperature (θ') fields at $t = 0$, 7 and 10 min from the benchmark run are shown in Fig. 7. Despite the different numerical techniques used, our results at 7 and 10 min compare remarkably well with those of Robert (1993) (who used a semi-Lagrangian semi-implicit method). It can be seen in Fig. 7 that sharp gradients close to two grid intervals are well preserved in the θ' field by the FCT advection. Hardly any undershooting or overshooting is found in the solution.

In the next four experiments (A4D4, A4MD4, A4D6, and A4MD6 in Table 1), a regular LT4CS scheme is used for the advection of both temperature and momentum. Such a scheme is known to introduce dispersive ripples or, in other words, overshoot and undershoot near sharp boundaries (see, e.g., Haltiner and Williams 1980). Our purpose here is to see if the monotonic option for high-order diffusion helps to reduce such over- and undershoot more effectively than a regular version of the diffusion.

Figure 8 shows the θ' field at $t = 7$ min from experiments using regular fourth-order (A4D4) and the simple monotonic fourth-order diffusion (A4MD4) for both temperature and momentum. Compared to Fig. 7b, both solutions are significantly smoothed, and more so for the regular fourth-order diffusion case. Overshooting is seen in both cases, with the one associated with monotonic diffusion being smaller (0.0318 vs 0.0656 K). The

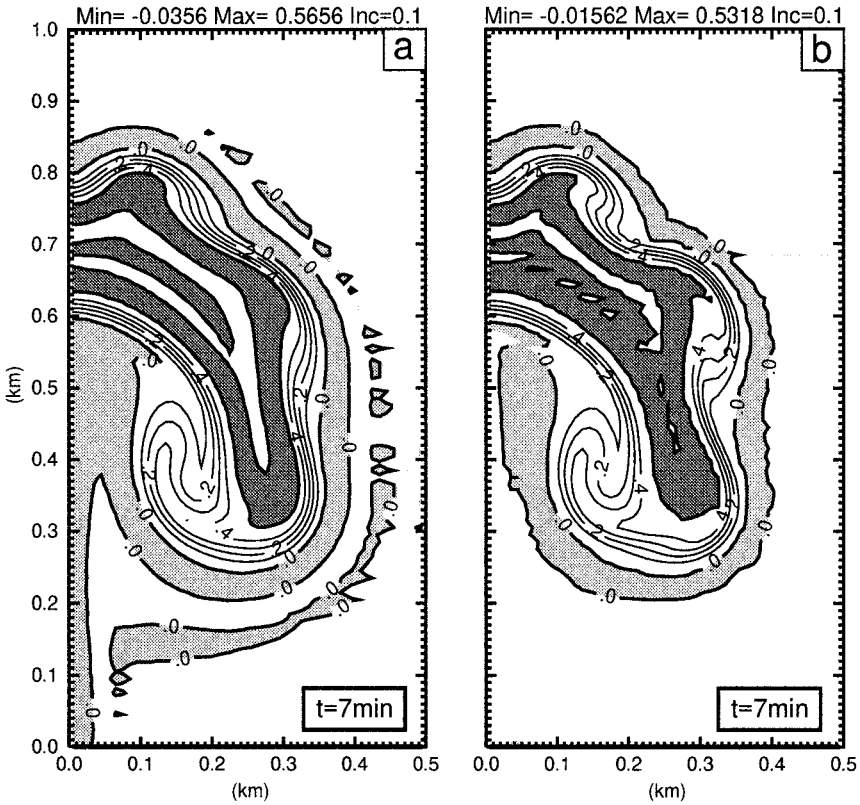

FIG. 8. Perturbation potential temperature fields at $t = 7$ min. Regular LT4CS scheme is used for advection of both temperature and momentum. In (a), regular fourth-order diffusion and in (b) simple monotonic fourth-order diffusion is applied to both temperature and momentum fields. Plotting conventions are the same as in Fig. 7. Overshooting and undershooting occur in both cases, as indicated by the minimum and maximum values in the plots. They are caused by both advection and diffusion in (a) and by advection alone in (b). Both undershooting and overshooting are smaller in the latter case where monotonic fourth-order diffusion is used. These plots should be compared with Fig. 7b.

additional overshooting and resultant smoothing can be attributed to the nonmonotonic behavior of diffusion, as illustrated by earlier 1D and 2D experiments. The maximum overshooting occurs inside the thermal plume near the top of the bubble, where the vertical velocity is largest (advection normal to the θ' gradient is strongest) and θ' field has a sharp plateaulike distribution (due to strong convergence at the front). While monotonic diffusion does not produce overshooting itself, it does not prevent the advection scheme from producing it. Therefore overshooting is still found in the solution, but less than the nonmonotonic case. Table 1 shows that both rms (root-mean-square) and maximum absolute errors (with respect to the benchmark) are smaller (0.0926 vs 0.1072 for the former and 0.5012 vs 0.527 for the latter) in the monotonic diffusion case (A4MD4) than the regular fourth-order diffusion case (A4D4).

Figure 9 shows results from two other experiments (A4D6 and A4MD6 in Table 1) in which the sixth-order diffusion schemes are used. Clearly the sixth-order diffusion schemes preserve sharp gradients much better than the fourth-order ones and the solutions here are much closer to the benchmark in Fig. 7. The latter is also supported by the rms error statistics in Table 1. Unfortunately the overshooting in θ' (0.0994 K in Fig. 9a and 0.0749 K in Fig. 9b) is also more serious. This is reflected also in the maximum absolute errors given in Table 1. The larger overshooting in A4MD6 compared to A4MD4 is attributed to advection in the presence of strong gradient made possible by the more selective sixth-order diffusion. The topmost thermal plume rose too fast as a result. Still, the solution using monotonic diffusion is better.

Considering that the sixth-order diffusion is more selective, one can afford a larger diffusion coefficient (as shown in Table 1, even the regular sixth-order diffusion performed better than the monotonic fourth-order diffusion in terms of rms error). We repeat the above two experiments with doubled diffusion coefficients (A4dD6 and A4dMD6); the maximum θ' values are reduced to 0.570 and 0.546 K, respectively. The maximum absolute errors are also reduced (to 0.5685 and 0.5318, respectively) but the rms errors are slightly increased but still less than their fourth-order counterparts. The θ' fields

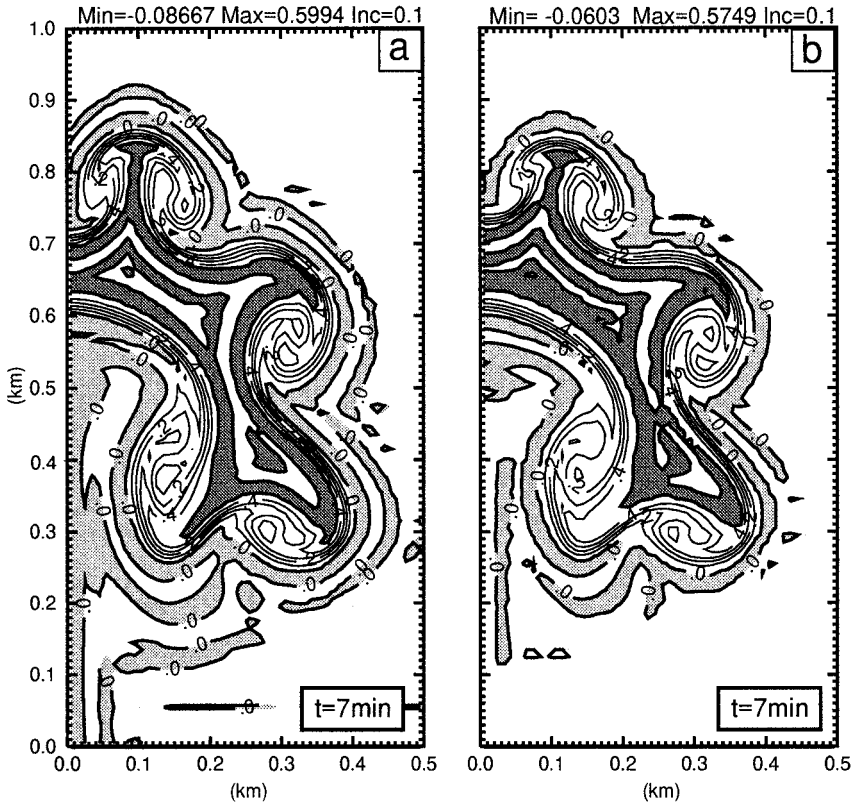

FIG. 9. As in Fig. 8, except in (a) regular sixth-order diffusion and in (b) simple monotonic sixth-order diffusion is used. Again, undershooting and overshooting are smaller in the (b) where monotonic sixth-order diffusion is used.

(not shown) are much closer to the benchmark solutions than the corresponding fourth-order ones.

In the benchmark case, no diffusion is applied to the θ field since no process in the run can generate ripples. In practice, when physical processes and external force are included, small-scale ($\sim 2\Delta x$) noise that are poorly handled by almost all numerical schemes can still be generated even if monotonic advection scheme is used. We conducted two more experiments that are the same as the benchmark run, except that additional sixth-order regular (FCTD6) or monotonic (FCTMD6) diffusion is applied to the θ' field. Since FCT advection preserves monotonicity, all undershooting and overshooting found are generated by the diffusion scheme alone or by the interaction of diffusion with advection. This therefore allows us to better gauge the effects of diffusion.

Figure 10 shows the results of these two runs. The undershoot (0.0554) and overshoot (0.0834) are rather large in the regular sixth-order diffusion case, supporting our finding in the simple tests with static square waves. Those in the monotonic diffusion case are much smaller (0.0032 and 0.0056) and the solution is very close to that of benchmark. The field appears slightly smoother, yet it still retains the sharp gradient at the edge of the bubble. Both rms and absolute errors are smallest among all experiments discussed so far (see Table 1). We attribute the small amount of overshooting and undershooting to the interaction between the advection and diffusion scheme. Perhaps true monotonicity can only be achieved by strictly enforcing it when the limiting of both advective and diffusive fluxes are considered together. This can be a topic of a future investigation while we remain focused on the general effectiveness of this diffusion scheme.

Finally, we present results from one more experiment (REF1). The experiment is the same as the benchmark except regular instead of monotonic sixth-order diffusion is applied to the momentum fields. The FCT advection scheme is used for θ' without any diffusion. Figure 11 shows the w fields at $t = 7$ min from this experiment (REF1) and the benchmark (REF). The topmost thermal plume is stronger and has risen higher due to larger w in REF1. This must have resulted from overshooting in the momentum fields caused by nonmonotonic sixth-order diffusion. Unlike with θ', we do not know the upper limit of w, it is therefore hard to determine the exact amount of overshoot in w. In the end, we comment that in the bubble experiments the off switch as given in Eq. (10) is typically active at less than 1% (around 0.4%) of the total number of grid points. This is of course flow dependent.

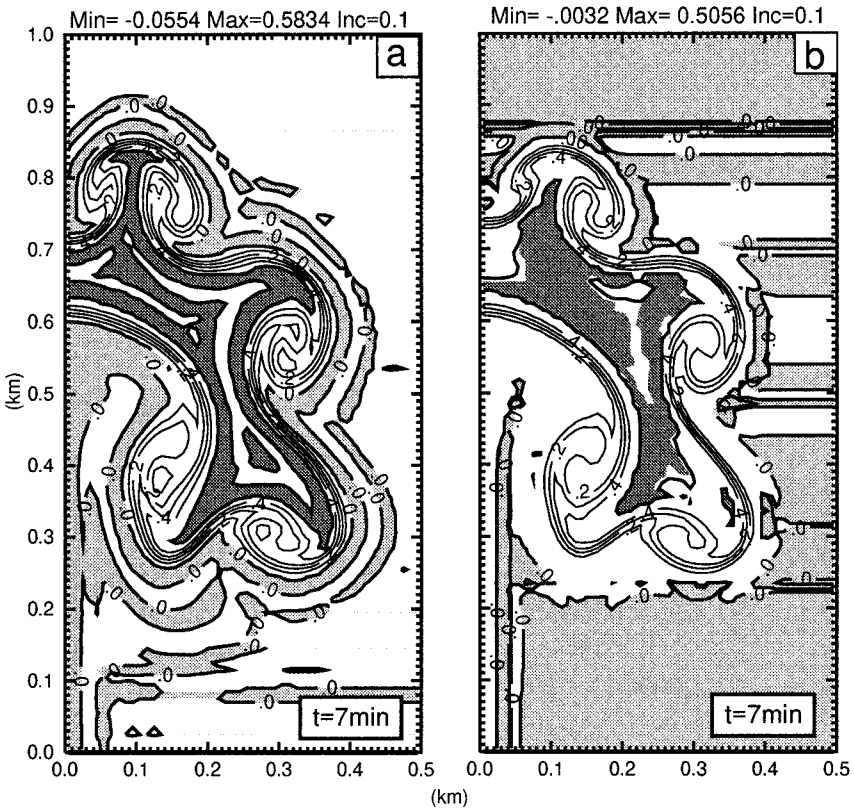

FIG. 10. As in Fig. 7 with the FCT scheme being used for potential temperature (θ) advection and simple monotonic sixth-order diffusion scheme used for momentum. In addition, (a) regular sixth-order diffusion and (b) simple monotonic sixth-order diffusion is also applied to θ. These plots should be compared with Fig. 7b where no diffusion is applied to θ. The solution in (b) is very close to Fig. 7b, and only very small over- and undershoot is found. Significant over- and undershoot exist in (a), causing the solution to differ from Fig. 7b.

6. Summary and discussion

In this paper, two types of monotonic high-order diffusion schemes are proposed and examined. The first is based on correction/limiting on the corrective flux, which is the difference between a high-order (fourth order and above) diffusion scheme and a lower-order (typically second order) one. Overshooting and undershooting found in the solutions of higher-order diffusions near sharp gradients are prevented, while the highly selective property of damping is retained. The flux limiter is general and can be applied to multidimensional diffusion terms.

An alternative and much simpler scheme is also proposed, which simply ensures that the diffusive fluxes are always downgradient; otherwise, the fluxes are set to zero. It is interesting that this much simpler scheme yields as good a solution in 1D cases as and better solutions in 2D (and 3D, not shown) than the one using a much more elaborate flux limiter. The scheme also preserves monotonicity in the solutions. In terms of computational cost, this scheme is much more attractive.

The simple flux-limited fourth- and sixth-order diffusion schemes are also applied to a thermal bubble convection case. It is shown that overshooting and undershooting are consistently smaller when the flux-limited version of the high-order diffusion is used, no matter whether the advection scheme is monotonic or not. This conclusion applies to both scalar and momentum fields. Higher-order monotonic diffusion works better and even more so when used together with monotonic advection.

We point out that step function–type distributions for which regular high-order diffusion schemes are most problematic are commonly found in numerical models and in the real world. Examples include the temperature field at the lower and upper boundaries of an inversion layer and at the edge of cold thunderstorm outflow, and in the liquid and ice water fields near the freezing level where liquid water turns into ice. At these sharp boundaries, most advection schemes will introduce small-scale noise that need to be controlled. When high-order diffusion is used to do so, it should not introduce new noise.

It should also be pointed out that the diffusive fluxes could be combined with the advective fluxes in a transport (advection) diffusion equation. Certain flux limiters can then be applied to the total flux so as to control overshooting and undershooting. However, it is not the purpose of this paper to study monotonic advection

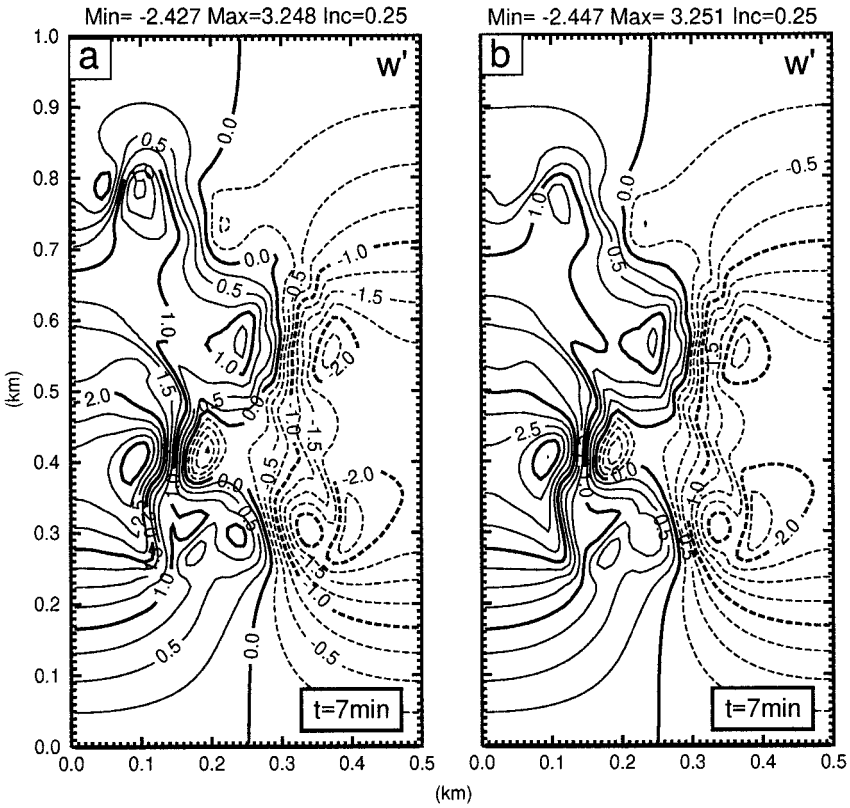

FIG. 11. The w fields at $t = 7$ min from (a) experiment REF1 and (b) benchmark REF. The topmost thermal plume is stronger and has risen higher in (a) due to larger w, caused by overshooting in momentum fields resulting from nonmonotonic diffusion.

schemes. In some cases, sophisticated monotonic advection scheme may be hard or even impossible to implement (e.g., for momentum equations in compressible models using mode splitting technique), they also tend to significantly increase the total cost and still may not completely remove the need of numerical diffusion (due to the presence of other processes). This is where our simple monotonic diffusion schemes fit in.

Acknowledgments. This work was supported by the National Science Function Grant ATM91-2009 to the Center for Analysis and Prediction of Storms, University of Oklahoma. Comments and suggestions on the draft by Drs. R. Carpenter, K. Droegemeier, and B. Fiedler improved this paper. The author also wants to thank two anonymous reviewers for their constructive suggestions.

REFERENCES

Boris, J. P., and D. L. Book, 1973: Flux-corrected transport. Part I: SHASTA, a fluid transport algorithm that works. *J. Comput. Phys.,* **11,** 38–69.

Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology*. John Wiley and Sons, 477 pp.

Hoskins, B. J., 1980: Representation of the earth topography using spherical harmonics. *Mon. Wea. Rev.,* **108,** 111–115.

Klemp, J. B., and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.,* **35,** 1070–1096.

Raymond, W. H., 1988: High-order low-pass implicit tangent filters for use in finite area calculations. *Mon. Wea. Rev.,* **116,** 2132–2141.

——, and A. Garder, 1976: Selective damping in a Galerkin method for solving wave problems with variable grids. *Mon. Wea. Rev.,* **104,** 1583–1590.

Robert, A., 1993: Bubble convection experiments with a semi-implicit formulation of the Euler equations. *J. Atmos. Sci.,* **50,** 1865–1873.

Sardeshmukh, P. D., and B. J. Hoskins, 1984: Spatial smoothing on the sphere. *Mon. Wea. Rev.,* **112,** 2524–2529.

Shapiro, R., 1970: Smoothing, filtering and boundary effects. *Rev. Geophys. Space Phys.,* **8,** 359–387.

——, 1975: Linear filtering. *Math. Comp.,* **29,** 1094–1097.

Skamarock, W. C., and J. B. Klemp, 1992: The stability of time-split numerical methods for the hydrostatic and nonhydrostatic elastic equations. *Mon. Wea. Rev.,* **120,** 2109–2127.

Smolarkiewicz, P. K., and W. W. Grabowski, 1990: The multidimensional positive definite advection transport algorithm: Nonoscillatory option. *J. Comput. Phys.,* **86,** 355–375.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, and K. Brewster, 1995: ARPS Version 4.0 User's Guide. Center for Analysis and Prediction of Storms, 380 pp. [Available from CAPS, University of Oklahoma, 100 E. Boyd St., Norman OK 73019.]

Zalesak, S. T., 1979: Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comput. Phys.,* **31,** 335–362.