The Eta slope adjustment: Contender for an optimal steepening in a piecewise-linear advection scheme? Comparison tests

By Fedor Mesinger*[1,2] and Dusan Jovic[2]

[1] NCEP Environmental Modeling Center, Camp Springs, Maryland

[2] The Abdus Salam International Centre for Theoretical Physics, Miramare, Trieste, Italy

17 October 2000, Revised 3 September 2002

NCEP Office Note

439

Summary

Attractive features of a fairly simple linear advection scheme are pointed out: the scheme is stable within the CFL condition, conservative, monotonic, and does not falsely amplify extrema during the advection process. It is a piecewise-linear scheme, with an iterative adjustment of slopes within the grid zones, aiming for what may be the most desirable shape within the piecewise-linear framework. Thus, in addition to the features listed, the scheme is "strictly monotonic", the term we suggest to denote that not only will monotonicity be preserved following a time step, but that the piecewise-linear representation prior to the actual transport within the time step is also monotonic. The scheme has been used for some time for the vertical advection of moisture variables in the Eta Model.

In standard top-hat advection tests the scheme is compared against five other schemes: an antidiffusion scheme of Janjic which is used for horizontal advection of moisture variables in the Eta; piecewise-linear schemes using three apparently leading slope definition algorithms: the so-called minmod, van Leer's (1974), or harmonic slope averaging, and the monotonized centered (MC) algorithm; and, finally, the Takacs' third-order scheme. The Eta slope-adjustment scheme, using three iterations, has in these tests performed significantly better than four of the test schemes, and, albeit with a slight margin, also better than the MC algorithm scheme. Additional tests suggest that this advantage over the MC scheme comes without sacrificing accuracy on smoother solutions.

Keywords: Advection Eta Model Finite-volume schemes Transport schemes

---

* Corresponding author: Environmental Modeling Center, 5200 Auth Road, Room 207, Camp Springs, MD 20746-4304. E-mail: fedor.mesinger@noaa.gov

1. Introduction

For a variety of good reasons, over the years the linear advection has kept its place as one of the standard topics of weather and climate modeling; just as it has in other fields of computational fluid dynamics. The ever increasing role and number of moisture variables in today's atmospheric models, with their "lumpy" character and sensitivity to the creation of false extrema, only adds to the importance of the problem. Even so, the use in weather and climate models of highly accurate and not all that complex finite-volume, "physics friendly" one might say schemes, and more specifically piecewise-polynomial schemes, is far from widespread. Perhaps the vast number of alternative possibilities -- also sporting a variety of attractive features -- along with the multitude of topics enjoying the attention of modelers, are the reasons. One example of the use of a finite-volume scheme in a GCM is that of Thuburn (1993): the so-called van Leer (van Leer 1974) flux-limited scheme was extensively tested in simple advection problems, and comprehensive GCM runs, and its positive impact stressed. Studies of Carpenter et al. (1990) of a still much more ambitious fully piecewise-parabolic method, and of Rancic (1992) of a semi-Lagrangian biparabolic scheme for passive scalars, are two of the rare examples of piecewise-polynomial efforts.

A piecewise-linear vertical advection scheme for moisture was implemented in the Eta Model some years ago (Mesinger 1988). This was motivated by a suspicion that centers of excessive precipitation ("bulls eyes") plaguing the model at the time might be partially caused by overshoots due to the centered vertical advection of moisture. Indeed, the replacement of the Lorenz-Arakawa centered scheme by the piecewise-linear advection visibly alleviated the problem (Mesinger and Janjic 1990).

We find it desirable to return to this topic for a number of reasons. One is the already referred to increased importance of the advection of moisture variables in forecasting models. Accurate transport of passive substances, such as airborne dust and chemical tracers, is also being frequently sought. Another of our incentives is the fact that a different scheme was implemented in the Eta subsequently for horizontal advection of moisture variables(Janjic 1997) so that a comparison of the performance of the two is an issue. Yet another is the vastly increased variety of approaches and schemes for linear advection that has become available since about the mid-eighties, so that tests of the performance of the Eta scheme against some of the other contenders we see as also desirable.

In section 2 we shall summarize some of the developments referred to and introduce definitions helpful for our later presentations. In section 3 we will define the iterative slope adjustment (SA) scheme used in the Eta. In section 4 we shall present results of the comparison tests performed against five other schemes: an antidiffusion scheme of Janjic (1997); piecewise-

linear schemes using three alternative slope definition algorithms: the so-called minmod, van Leer's (1974), and the monotonized centered (MC) algorithm; and, finally, the Takacs' third-order "minimized dissipation and dispersion errors" scheme. Aiming to illuminate the background for the results obtained, in section 5 we shall look into the behavior of the SA scheme for simple distributions of interest, and will present tests on the dependence of result on the number of iterations performed. We shall end with a few concluding comments.

2. Background and definitions

A number of developments during the seventies and the early eighties have very much invigorated the field of the linear advection or transport schemes, addressing with considerable success the classic difficulties of the creation of false extrema and parasitic waves (e.g., Wurtele 1961) during the advection process. One that very quickly found its application in meteorology is that of the advent of *flux-corrected transport*, or FCT schemes, introduced by Boris and Book (1973). With the FCT approach, writing the advection equation in flux (or "conservation law") form preliminary fluxes are corrected aiming to remove their diffusive effect but yet so as not to create new maxima or minima as a result. Since the corrective fluxes are commonly directed up-gradient, the term *antidiffusive* fluxes is used for these corrective, second-step fluxes.

A similar strategy is that of not actually computing the preliminary fluxes, but of apportioning the fluxes to be used between the typically simple upstream preliminary fluxes, and some chosen higher-order ones. Since this apportioning is customarily expressed via a function multiplying the difference between the two, *flux-limiter methods* is the term used for the class of schemes obtained.

For definitions of several test schemes that we shall use, an illustration of this formalism will be useful. To this end, consider the simplest case of the one-dimensional linear advection equation

$$\frac{\partial q}{\partial t} + c\frac{\partial q}{\partial x} = 0, \quad c = const > 0 \ , \tag{1}$$

with symbols here and further on, unless specifically defined, having their usual meaning. (1), of course, is equivalent to its flux form

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}(cq) = 0 \ . \tag{2}$$

Discretizing (2), we assume grid intervals of equal length, and -- for a finite-volume

approach -- take grid point values q_j to represent averages of $q(x)$ over the grid intervals. The *conservation form* approximation to (2) then is

$$q_j^{n+1} = q_j^n - \frac{\Delta t}{\Delta x}(F_{j+1/2} - F_{j-1/2}) \ , \tag{3}$$

with $F_{j+1/2}$ and $F_{j-1/2}$ being analogs of fluxes of q , cq, at grid mesh boundaries, used to advance the solution according to a time-differencing algorithm yet to be decided upon.

To accommodate a later representation of $q(x,t)$, associated with $F_{j+1/2}$ and $F_{j-1/2}$, in form of a function which is continuous and more specifically linear inside each grid interval but possibly discontinuous at mesh boundaries, we denote

$$F_{j+1/2} \equiv c\,(q_{j+1/2})^- \ , \quad F_{j-1/2} \equiv c\,(q_{j-1/2})^- \ , \tag{4}$$

the superscripts - indicating that the approximations of q at $(j+1/2)\Delta x$ and at $(j-1/2)\Delta x$ used are valid at the sides of the smaller values of x. This, of course, is a choice due to our considering the case of positive c; + superscripts would be in place if c were negative. The simplest illustration of the outlined framework then is

$$(q_{j+1/2})^- = q_j + \frac{1}{2}(q_{j+1} - q_j)\, C \ , \tag{5}$$

with C here being the flux *limiter*. Note that if for the analogs of fluxes in (3) a forward-in-time scheme were used, and if C were equal to 0, (3) through (5) would represent the simple upstream scheme; and if it were equal to 1, it would be the forward-in-time space-centered scheme. More refined strategies are actually used in flux-limiter methods, and also more complex choices of C. We shall refer to some choices of C in the following section, as they are applicable not only to the flux-limited, but also to piecewise-linear, or *slope*-limited schemes, that we shall be testing.

We now turn our attention to piecewise-polynomial, and specifically piecewise-linear schemes. Rather than using a difference analog such as (3), with these schemes the procedure for advancing the numerical solution of (2) one time step ahead is equivalent to the three steps as follows. In the first step, a piecewise-linear distribution of $q(x,t^n)$ is defined, linear inside each grid mesh, that is

$$q(x,t^n) = q_j^n + \sigma_j^n(x - x_j) \qquad \text{for} \quad x_{j-1/2} \le x \le x_{j+1/2} \ . \tag{6}$$

In the second step, the advection equation is solved exactly over the time step Δt. With our

linear representation, (6), this is equivalent to using fluxes valid at the middle of the time step, as opposed to the forward-in-time values (4). Finally, in the third step, the resulting $q(x,t^{n+1})$ is averaged over the grid intervals to arrive at new grid point values q_j^{n+1} . In practice, the second and the third step need not be done separately, as, given (6), changes in grid point values are easily calculated directly as a result of transports through the two mesh boundaries.

Note that it is this exact integration of the advection equation in the second step that makes the piecewise-linear schemes' approach different from that of the flux-limited methods. Namely, integrals of fluxes over the time step are used to update the grid point values, as opposed to the instantaneous fluxes, (4), which would be the case with the flux-limited schemes. However, (5) can just as well be used to define the slopes inside grid zones, and in fact it commonly is; C is then referred to as a *slope*-limiter (e.g., LeVeque 1996, Durran 1999).

3. The Eta slope-adjustment scheme

Given a distribution of grid point values q_j^n , q_j for brevity, a definition of slopes is required by (6). Our starting point is that of a piecewise-constant distribution, that is,

$$\sigma_j = 0 \tag{7}$$

at all grid points j. We intend to arrive at final slopes to be used in (6) in an iterative way; thus, let another subscript, i, denote the number of iterations done. (7) can then be written as

$$(\sigma_j)_i = 0 \qquad \text{for} \quad i = 0 \,. \tag{8}$$

The standard way of guaranteeing no false amplification of extrema is to keep the slopes zero at extremum points. In general discontinuities will exist at mesh boundaries at points in between neighboring extrema. It appears to us desirable from the physical point of view to construct slopes so as to steepen the slopes in between the extrema to have these discontinuities removed or reduced on an equal basis by the steepening -- if it is to be performed -- at neighboring zones, but only to the extent this is possible without creating new discontinuities of an opposite sign. We refer to this requirement as one of maintaining "strict" monotonicity: not only will monotonicity be preserved once the integration is done for the time step, but also it will not be violated by the relation (6) of step one of the scheme.

The discontinuity at a mesh boundary possibly available for the slope adjustment within the iteration i is

$$(\delta q_{j+1/2})_i = (q_{j+1/2})^+_{i-1} - (q_{j+1/2})^-_{i-1} \ , \tag{9}$$

positive in an increasing segment of the distribution, that is, for $q_{j+1} > q_j$. At a grid point which is not an extremum two of these need to be considered, one at each mesh boundary. To maintain strict monotonicity as defined, only half of the smaller one, in absolute value, can safely be consumed in a slope adjustment iteration unless a neighbor is an extremum in which case all of the discontinuity on its side is available.

This is illustrated in Fig 1. A starting distribution over five zones j+1 through j+5 is shown by the dashed line, with a minimum at j+1, and a maximum at j+5. Thus, the slopes at these two zones will remain zero. At the four boundaries in between these zones the discontinuities (9) in the example depicted initially are, in relative units, 2, 8, 4, and 1. At the two boundaries next to the extrema all of the respective discontinuities, or steps, 2 and 1 in the example shown, are available for the construction of slopes in zones neighboring the extrema; at the two remaining boundaries, only half are available. For strict monotonicity, at each zone for which the slope is allowed the smaller of the "available steps" at its two boundaries is used to define the slopes; if the grid distance is chosen for unit of length in the example shown resulting in first iteration slopes of 2, 2, and 1, shown by the solid lines. Discontinuities remaining after the first iteration, 4 and 1 in the figure, can be reduced by subsequent iterations.

In other words, we define our slope iterations by

$$(\sigma_j)_i = (\sigma_j)_{i-1} + (\Delta\sigma_j)_i \ , \quad i = 1, 2, ..., N, \tag{10}$$

where

$$(\Delta\sigma_j)_i = \frac{a_j}{\Delta x}\left(\text{sgn}\,[(\Delta q)_{j+1/2}] \min\,[|(1-\tfrac{1}{2}a_{j-1})(\delta q_{j-1/2})_i|, |(1-\tfrac{1}{2}a_{j+1})(\delta q_{j+1/2})_i|] \right) \ , \tag{11}$$

with

$$(\Delta q)_{j+1/2} = q_{j+1} - q_j \ ,$$

and

$$a_j = \frac{1}{2}\left(1 + \frac{(\Delta q)_{j-1/2}(\Delta q)_{j+1/2}}{\left|(\Delta q)_{j-1/2}(\Delta q)_{j+1/2}\right|}\right), \quad \text{if} \quad (\Delta q)_{j-1/2}(\Delta q)_{j+1/2} \neq 0 \ ,$$

$$a_j = 0 \ , \quad \text{if} \quad (\Delta q)_{j-1/2}(\Delta q)_{j+1/2} = 0 \ . \tag{12}$$

a_j, thus, denotes the allowance of the slope adjustment in a zone, with 1 and 0 denoting that the

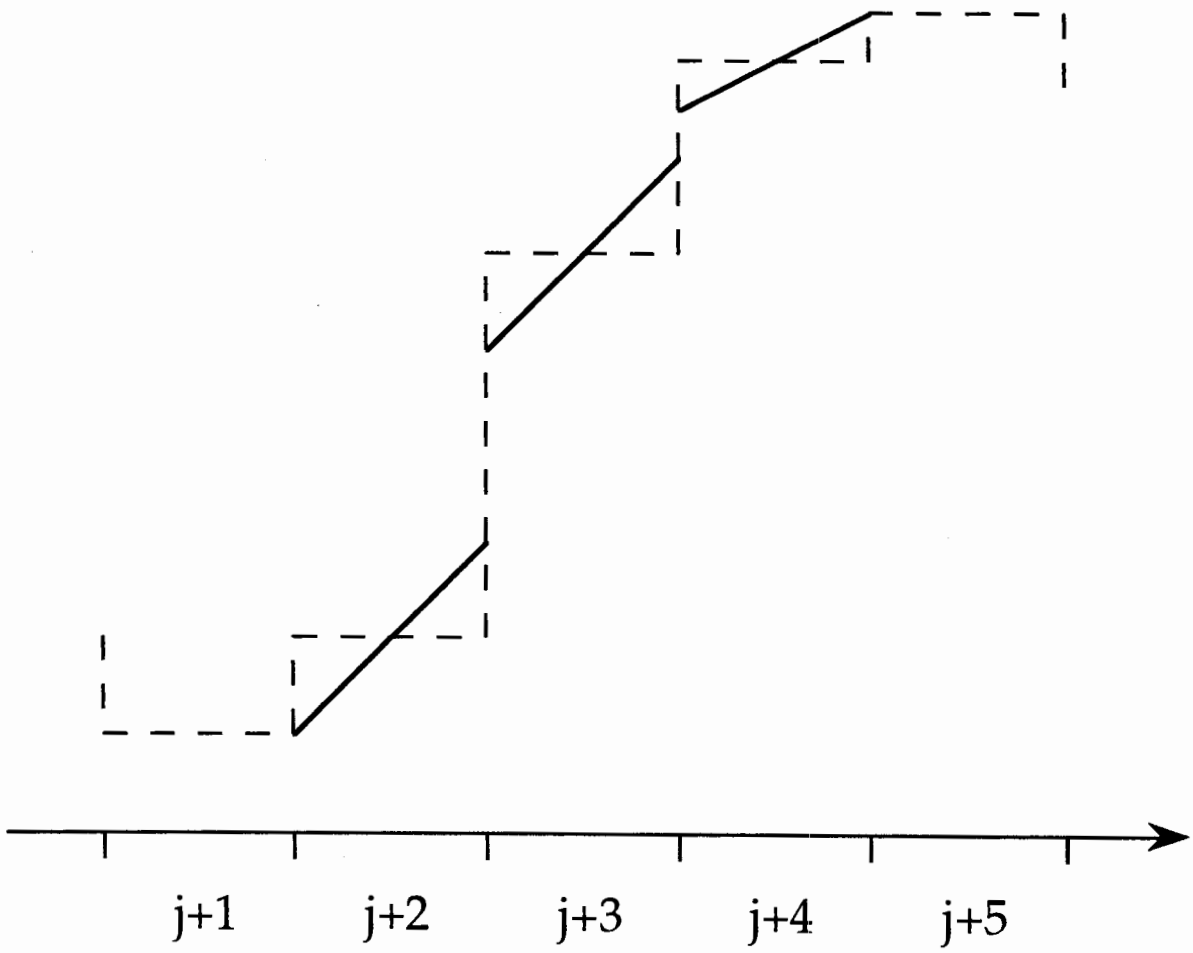

Figure 1. An example of the Eta iterative slope adjustment algorithm. The initial distribution is illustrated by the dashed line, with slopes in all five zones shown equal to zero. Slopes resulting from the first iteration are shown by the solid lines. See text for additional detail.

slope adjustment is allowed or not allowed, respectively. Three iterations are performed in NCEP's operational version of the Eta.

4. Comparison tests

We will here display and comment on results obtained using five schemes we have chosen for our comparison tests. The first on our list is the scheme used for horizontal advection of moisture variables in the Eta: with both schemes used in the model, it looked like an obvious candidate. The scheme, due to Janjic (1997), perhaps can also serve as a prototype for a maximally simplified antidiffusion scheme. The next three schemes are piecewise linear schemes, using three slope limiters of the four that both LeVeque (1996) and Durran (1999) have on their short lists of "standard" and "possible" limiters, respectively. The fourth limiter on these lists, by its inventor nicknamed superbee (Roe 1985), has been held in a somewhat lower esteem by the cited authors, for tending to be "overcompressive" (LeVeque), and for producing steepnesses greater than those of the true solution near the extrema of a sine wave (Durran). The fifth scheme on our list is the third-order scheme of Takacs (1985). Note that the Takacs' scheme has been implemented for both horizontal and vertical advection of all variables in a nonhydrostatic version of the Eta of Gallus and Rancic (1996).

For all five schemes tests were done for a top-hat function, spanning over 20 grid-intervals of a 100-interval one-dimensional domain. Tests were run for two translations across the domain, and for Courant numbers at 0.1 intervals running from 0.1 to 0.9. Definitions of schemes will be given, and plots for Courant numbers of 0.1 and 0.7 will be shown. For the monotonized-centered (MC) limiter, subsection (d), tests were done and are presented for two additional distributions, one consisting of a sine and a pulse wave, and the other for a narrow square wave, a top-hat spanning over 3 grid intervals only of a 13-interval domain. The purpose of this last test was to achieve a better focus on the difference in results arrived at by the two schemes, in view of the small difference in results of the 20 grid-intervals top-hat test.

(a) *The Janjic antidiffusion scheme*

With this scheme (Janjic 1997), as in Smolarkiewicz (1983), the first step is done upstream and forward in time, thus,

$$q_j^{n+1\,*} = q_j^n - \mu\,(q_j^n - q_{j-1}^n)\ , \tag{13}$$

where

$$\mu \equiv c\,\Delta t\,/\,\Delta x$$

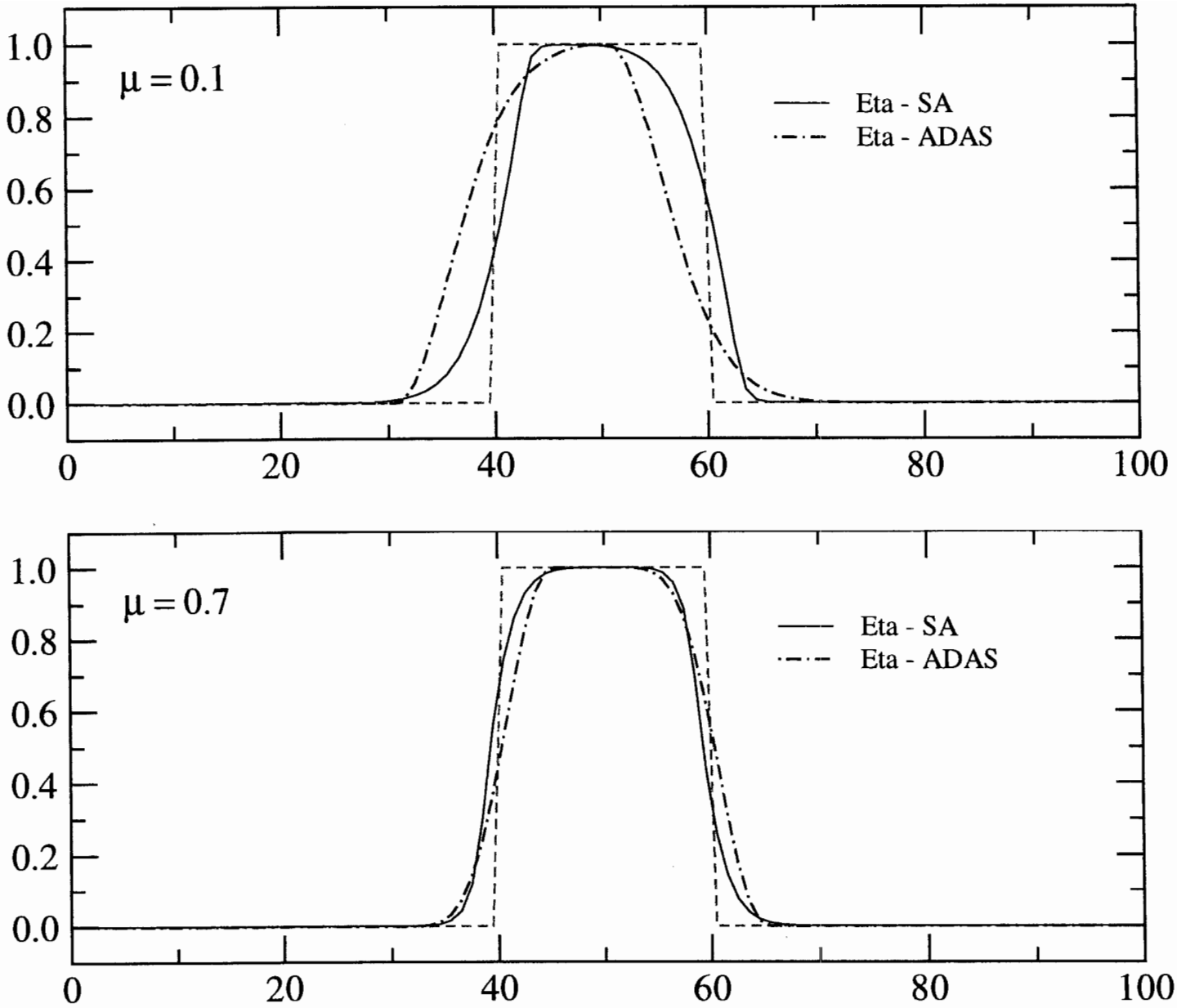

Figure 2. Results of top-hat advection tests using the Eta piecewise-linear slope-adjustment scheme (SA, solid line), and using the Eta (Janjic 1997) antidiffusion scheme (ADAS, dot-dashed line), after two translations of the true solution across the domain. The true solution is shown by the dashed line. Top panel shows results for the Courant number of 0.1, and lower panel those for the Courant number of 0.7, requiring 2000 and 286 time-steps, respectively.

is the Courant number. The second, antidiffusion step, is however simplified to

$$q_j^{n+1} = q_j^{n+1*} - f(\mu)\,(q_{j-1}^{n+1*} - 2q_j^{n+1*} + q_{j+1}^{n+1*}) \ , \tag{14}$$

where the function $f(\mu)$ has been empirically optimized by Janjic, as

$$f(\mu) = 0.52500\,\mu - 0.64813\,\mu^2 + 0.24520\,\mu^3 - 0.12189\,\mu^4 \ . \tag{15}$$

Should however (14) turn out to be greater than both the q_j^n and q_{j-1}^n, it is replaced by the greater of these two. Conversely, should it be less than both of them, it is replaced by the smaller of the two. Subsequently, before actually updating the values of q_j, conservation is enforced on the obtained increments of q_j. This is done by comparing the sum of all the positive and all the negative increments, and by multiplying increments whose sum is smaller in absolute value by the ratio of the two, in such a way as to make them equal. Finally, having updated the values of q_j, since the conservation "fixer" (the term of Rasch and Williamson 1990) could have created negative values in case the negative increments were the ones magnified positive definiteness is enforced by filling.

Note that, as a result of the sequence and nature of these three fixers, positive definiteness will be strictly enforced, while prevention of the creation of false extrema and conservation will not. Violations of the last two are however expected to be extremely small.

Solutions obtained after two translations using the Eta piecewise-linear scheme of the preceding section (Eta - SA), and that of the described antidiffusion scheme (Eta - ADAS), are shown in Fig. 2 by the solid and by the dot-dashed line, respectively. The upper panel shows the solution for the Courant number of 0.1, that is, after 2000 time steps; and the lower panel shows the solutions for the Courant number closest to 0.7 with an integer number of time steps needed for one translation, which is 143. The true solution is shown by the dashed line.

A better success of the slope adjustment (SA, or SA3) scheme than that of the Eta antidiffusion scheme in resisting the erosion of the top-hat shape is obvious in the lower plot. In the upper plot it is somewhat obscured by the opposite sign of the phase speed error of the two. Therefore, for more quantitative information on the error in this case, as well as on errors for other Courant numbers, we have calculated the "total" rms error of the two solutions

$$E_{TOT} = \frac{1}{M}\,\Sigma\,(q_j - q_T)^2 \ , \tag{16}$$

q_T here standing for the true solution, and M for the total number of grid points. We have also,

following Takacs (1985), calculated the "dissipation" and the "dispersion" components of the total error

$$E_{TOT} = E_{DISS} + E_{DISP} \ , \tag{17}$$

where

$$E_{DISS} = \left(\sigma(q_T) - \sigma(q_j)\right)^2 + \left(\overline{q_T} - \overline{q_j}\right)^2 \ , \tag{18}$$

and

$$E_{DISP} = 2(1-\rho)\,\sigma(q_T)\,\sigma(q_j) \ . \tag{19}$$

Here, σ denotes standard deviation of a variable, ρ the correlation coefficient of q_j and q_T, and overbar the domain average. Note that (19) is zero for a perfect correlation of q_j and q_T, which explains its name. The errors of the two solutions, dissipation, dispersion and total, as functions of the Courant number, are shown in the three panels of Fig. 3.

The dispersion error, as one might have expected, is seen to dominate the total error. As to the errors of the upper plot of Fig. 2, as well as those for all the Courant numbers tested up to and including 0.6, a considerable advantage of the SA scheme is revealed.

(*b*) *Piecewise-linear scheme, minmod limiter*

The standard way of defining the slopes of piecewise-linear functions – followed by numerous authors – is that of taking the limiter C of (5) to be a function of the ratios of differences at j - 1/2 and at j + 1/2, that is,

$$C = C(r), \quad r \equiv \frac{q_j - q_{j-1}}{q_{j+1} - q_j} \ . \tag{20}$$

The "minmod" limiter (e.g., Durran 1999; originally used by Roe, and later by various other authors, Sweby 1985) is

$$C(r) = \max\,[0, \min\,(1,r)] \ . \tag{21}$$

An example of its effect is given in Fig. 5.16 of Durran (1999): it defines the slope to be that of the smaller, in absolute value, of the two boundary values of $\Delta q / \Delta x$ unless q_j is an extremum in which case the slope is zero.

Results obtained with the piecewise-linear scheme using the minmod limiter, along with

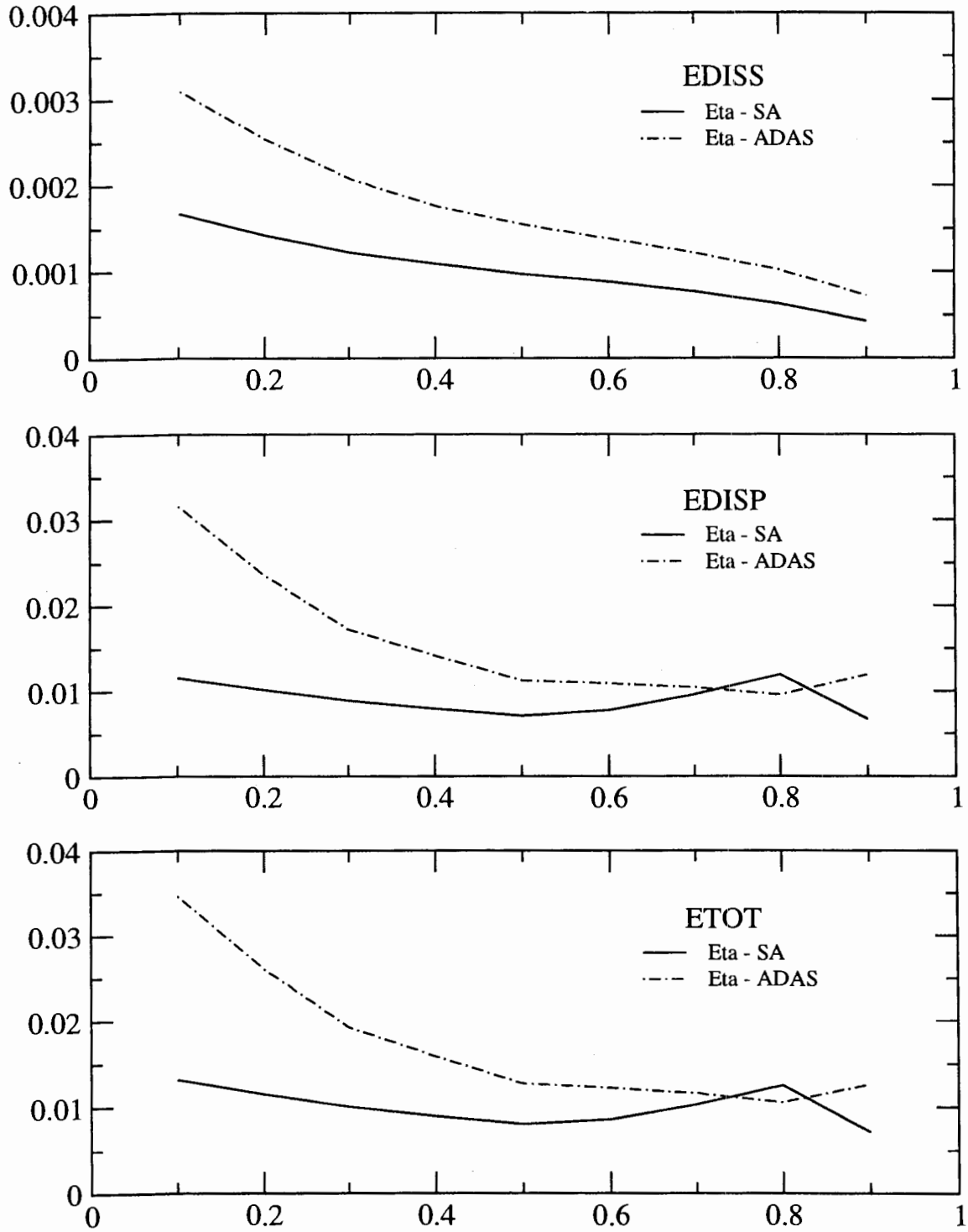

Figure 3. Dissipation, dispersion and total error of the slope-adjustment and the Eta antidiffusion scheme, as functions of Courant number, after two translations of the top-hat initial condition shown in Fig. 1 across the domain. See text for definitions of the errors.

those of the SA scheme, are shown in Fig. 4. Compared to the SA scheme, considerable erosion is seen to occur with the minmod limiter, no doubt due to its less-than-aggressive nature. The limiter does however distinguish itself, just as does the SA scheme, by preserving monotonicity within the intermediate first step of the scheme, that is, creating no sawtooth features in between monotonic segments of q_j ; a feature which we have named strict monotonicity. The remaining three limiters on the mentioned short lists do not share that property; an example of a violation of strict monotonicity, by the superbee limiter, will be illustrated later, in Fig. 11.

(c) *Van Leer limiter*

The Van Leer limiter (Van Leer 1974) is

$$C(r) = \frac{r + |r|}{1 + |r|} \,. \tag{22}$$

It results in a harmonic average of the two boundary slopes, unless q_j is an extremum in which case the slope is again zero. The limiter is monotonic; as pointed out by Van Leer, in order to preserve monotonicity, the linear function (6) must not take values outside the range spanned by the neighboring mesh averages. Recall that monotonicity is a property of the scheme not to create new extrema within a monotonic segment of q_j during the advection process. The same condition, applied at zones next to those of minima and maxima, will ensure no spurious amplification of the existing extrema during the advection. (22) is stronger than the referred to Van Leer's monotonicity condition, that is, it limits the slopes more than it is necessary for monotonicity. Van Leer limiter was used by Thuburn (1993) for his flux-limited scheme.

The results of our test for the piecewise-linear advection using the Van Leer limiter, again along with those for the SA scheme, are shown in Fig. 5. The Van Leer limiter scheme is seen to have performed not much worse than the SA scheme in preserving the top-hat shape; and much better than the minmod limiter scheme. Yet, the plots do illustrate a clear advantage of the SA scheme.

(d) *Monotonized-centered limiter*

The monotonized-centered (MC) limiter, also due to Van Leer (1977), is

$$C(r) = \max\left(0, \min\left(2r, \frac{1+r}{2}, 2\right)\right) . \tag{23}$$

It defines slopes to be the algebraic average of the two boundary slopes, that is, to be prescribed

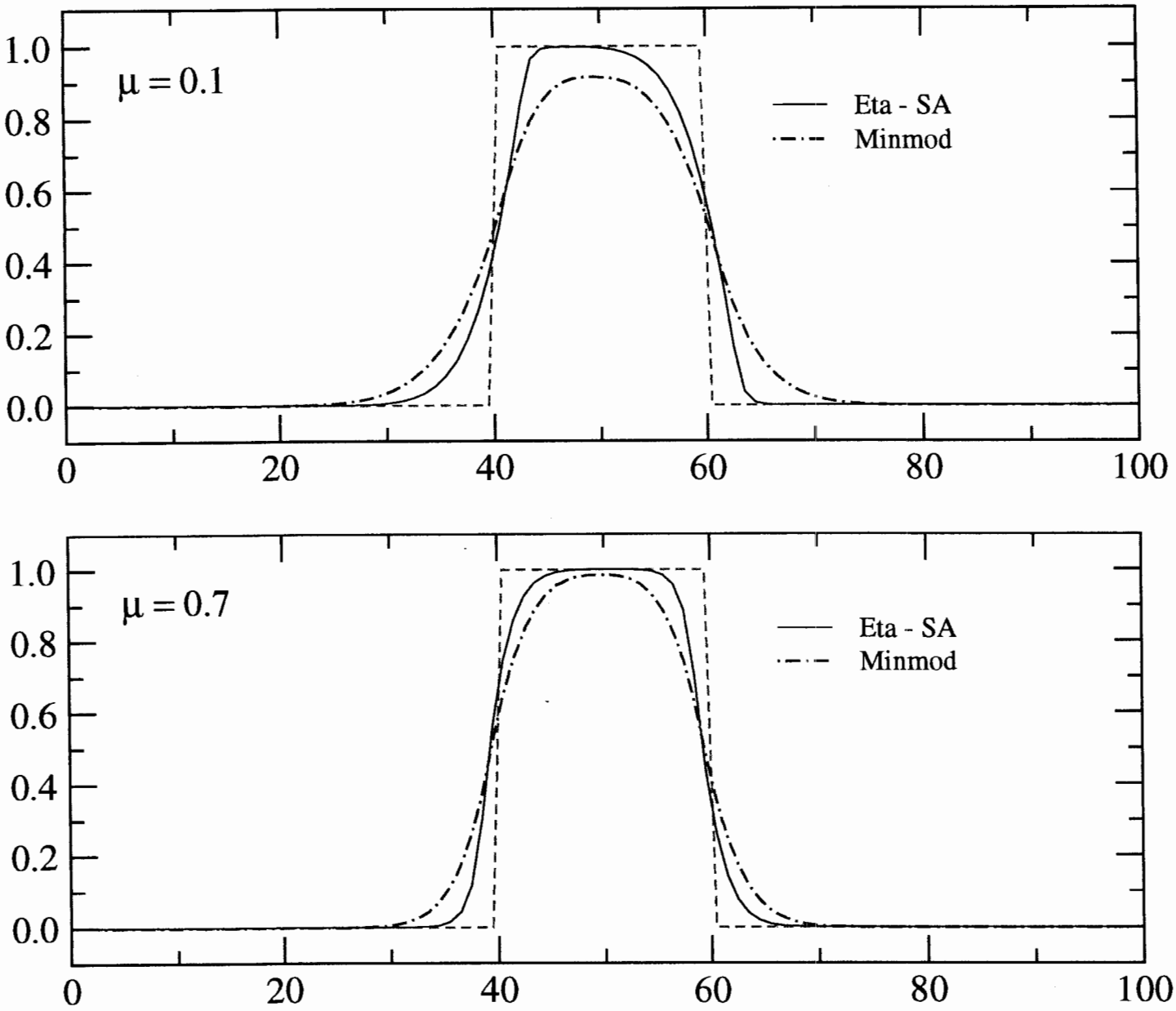

Figure 4. Same as Fig. 2, except for the Eta slope-adjustment scheme results (SA, solid line) compared against those using the minmod slope limiter (dot-dashed line). See text for definitions of schemes.

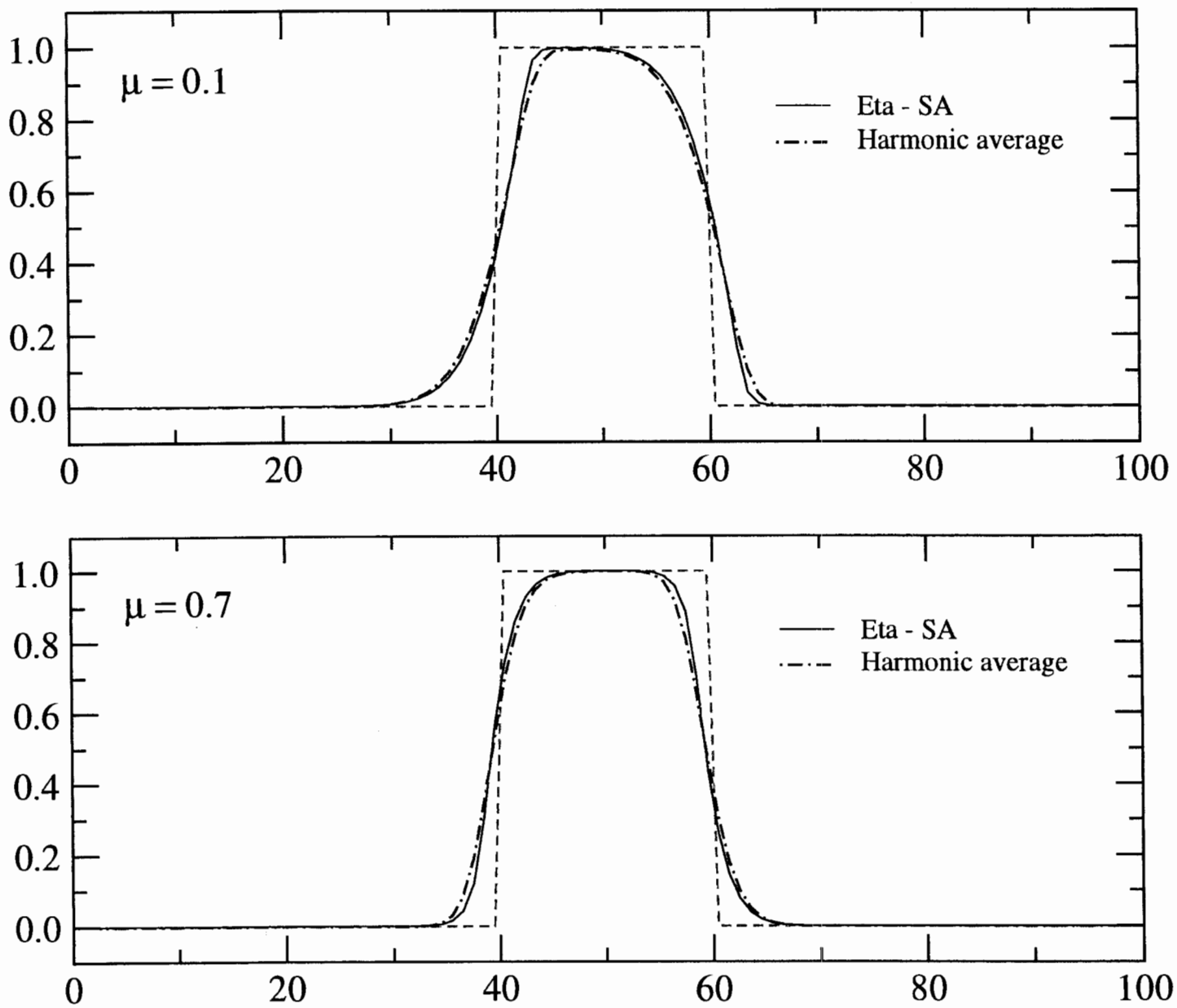

Figure 5. Same as Fig. 2, except for the Eta slope-adjustment scheme results (SA, solid line) compared against those using the Van Leer, or harmonic average, slope limiter (dot-dashed line). See text for definitions of schemes.

by a centered scheme; unless this violates the monotonicity condition in which case they are reduced to the extent required. If, however, q_j is an extremum the slope is once again set to zero.

Results of the test using the MC slope limiter, once more along with those for the SA scheme, are shown in Fig. 6. The two schemes are seen to have performed quite similarly. For both Courant numbers the SA scheme did a little better in maintaining the steepness of the slopes on both sides of the wave. However, for the Courant number of 0.1, this increased steepness of the SA scheme result came at a cost of its positive phase speed error being slightly greater than that of the MC scheme.

The comparison of the two this time being a close call, we have performed two additional tests. The first one, with the results shown in Fig. 7, was done for a sine and a pulse wave, spanning 40 intervals each. The difference in the two solutions, following two translations, is once again seen to be very small, this time with a slight advantage in favor of the MC scheme seen in its little better preservation of the amplitude of the sine wave for the Courant number of 0.1. But the more conspicuous errors, of shape modifications, and of the phase error for the Courant number of 0.7, for any practical purposes seem just about the same.

The other additional test we have performed for case of a narrow square wave, spanning 3 intervals of a 13-interval domain. We have done this test for the Courant number of 0.5; note that for the Courant number of 0.5 one can see by inspection that symmetry of the solution ought to be preserved. The results, following two translations, are shown in Fig. 8. An advantage of the SA over the MC scheme is now seen very clearly: the SA scheme values are higher than MC ones inside the analytic solution, and lower outside, in all zones that have values significantly greater than zero. We thus find that the SA scheme does appear to do significantly better on the square wave case with little if any penalty relative to the MC scheme on the smoother solutions.

(*e*) *The Takacs third-order scheme*

Takacs (1985) has considered the family of all two-level schemes in which the value of q_j^{n+1} is calculated as a linear combination of four grid-point values at time level n, ranging from j - 2 to j + 1. The scheme (13) and (14) without the fixers, incidentally, is a member of this family. The Takacs' family of two-level four-point schemes contained one third-order scheme. It is

$$\begin{aligned} q_j^{n+1} = q_j^n - \frac{\mu}{2}(q_{j+1}^n - q_{j-1}^n) + \frac{\mu^2}{2}(q_{j+1}^n - 2q_j^n + q_{j-1}^n) \\ + \frac{1}{6}\mu(1-\mu^2)(q_{j+1}^n - 3q_j^n + 3q_{j-1}^n - q_{j-2}^n)\,. \end{aligned} \tag{24}$$

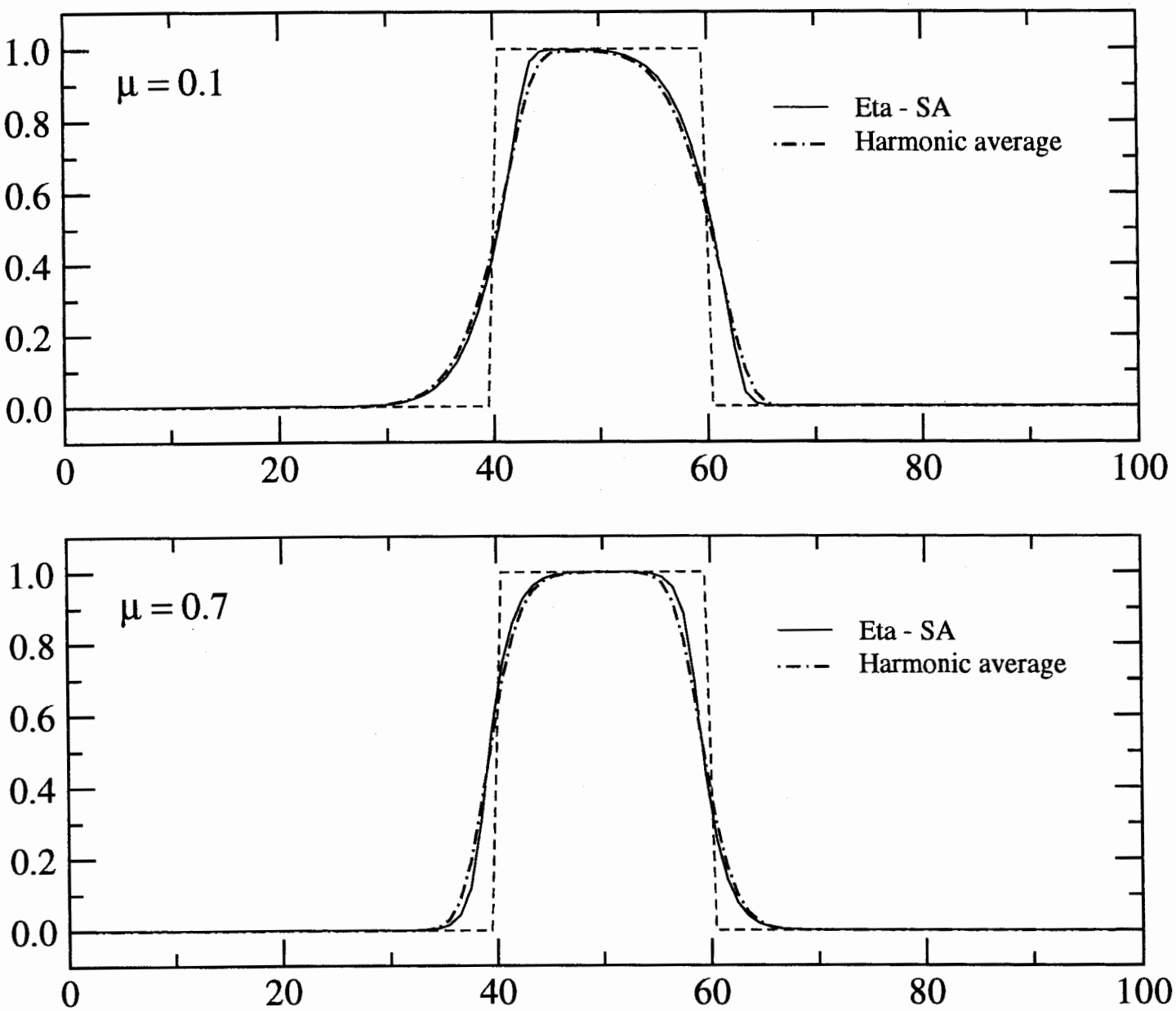

Figure 5. Same as Fig. 2, except for the Eta slope-adjustment scheme results (SA, solid line) compared against those using the Van Leer, or harmonic average, slope limiter (dot-dashed line). See text for definitions of schemes.

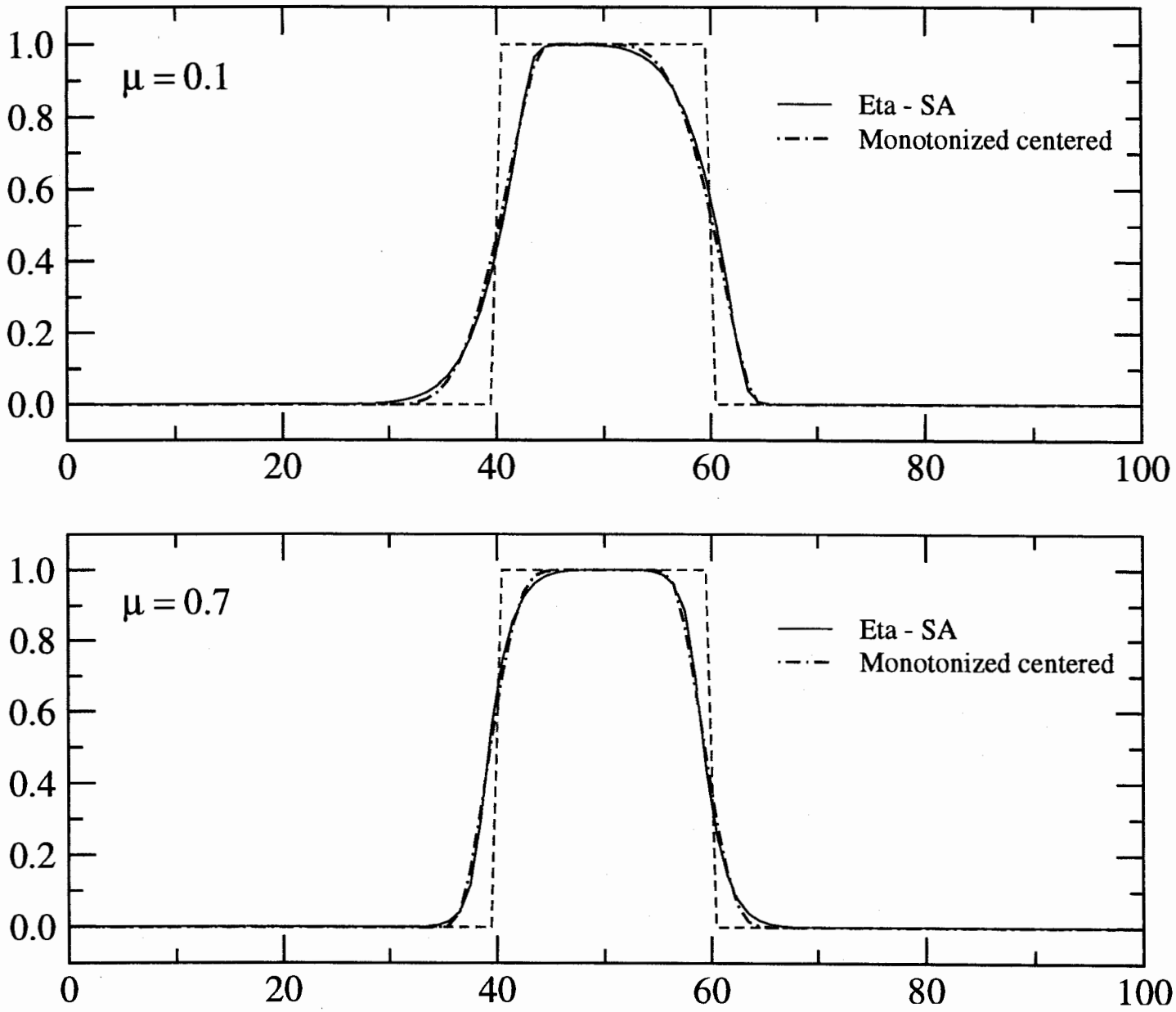

Figure 6. Same as Fig. 2, except for the Eta slope-adjustment scheme results (SA, solid line) compared against those using the monotonized centered slope limiter (dot-dashed line). See text for definitions of schemes.

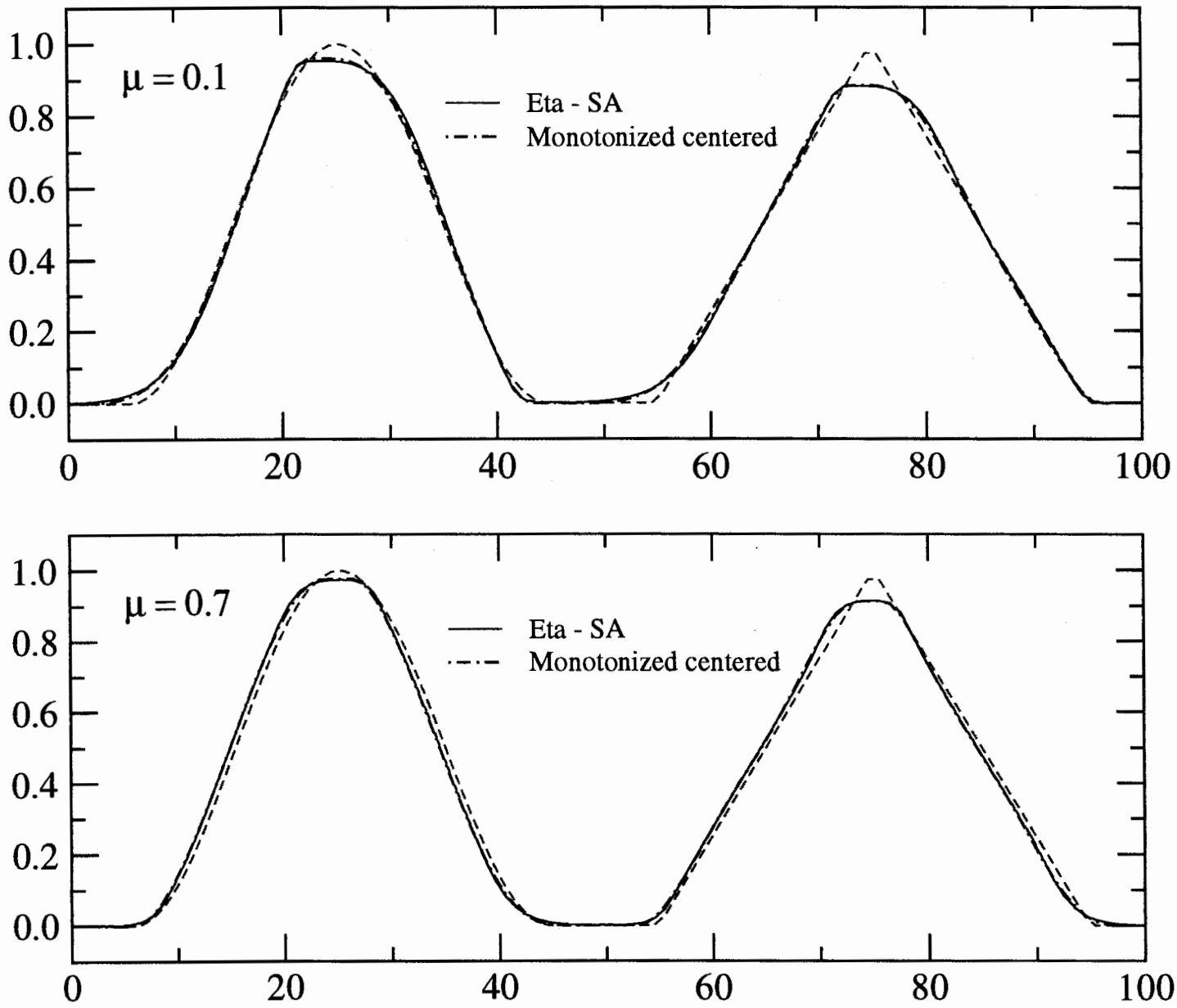

Figure 7. Same as Fig. 6, except for a sine plus a pulse wave initial distribution, spanning 40 zones each.

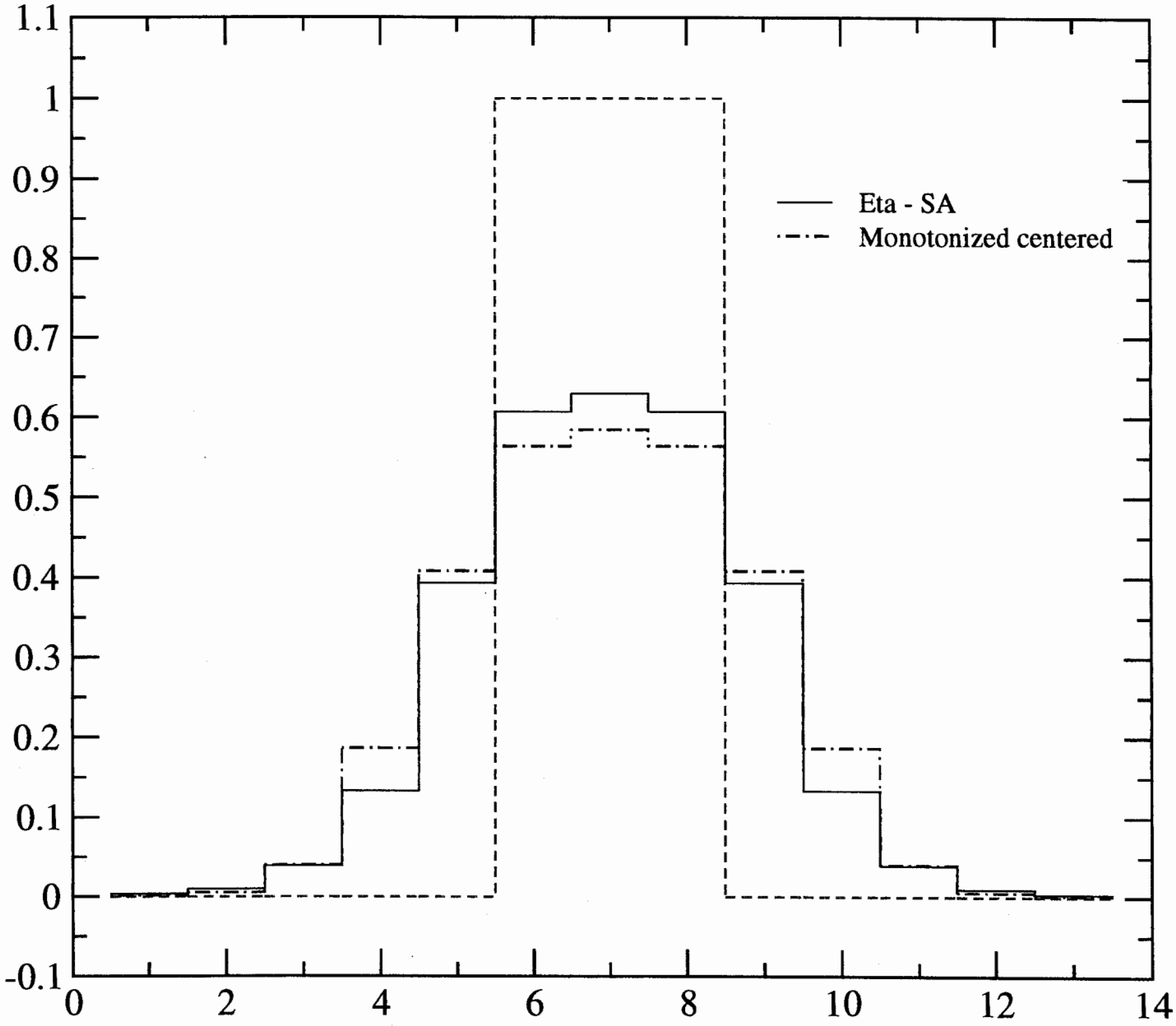

Figure 8. Same as Fig. 6, except for a "narrow" square wave initial distribution, spanning 3 zones, and a 13 zone domain.

Takacs' tests of this third-order scheme against the remaining stable but second-order members of the family led him to accord to the third-order scheme the attractive adjectives of two-step scheme "with minimized dissipation and dispersion errors".

Our tests using (24), presented in the same way as tests of the preceding 20-zone square-wave plots, are shown in Fig. 9. While the Takacs' scheme for the Courant number of 0.1 could be declared to have done slightly better than the SA scheme regarding the phase speed of the wave, its overshoots and undershoots do look unattractive. Given the variety of possible problems with conservation and other fixers (e.g., Rasch and Williamson 1990) that just about or in most cases would have to be used, as well as their largely ad-hoc nature, we feel that the overall advantage of the SA scheme in the results of Fig. 9 is hardly questionable.

5. Additional discussion and some remaining issues

We here raise several points that come to mind. As the first, one might wish to have a better feeling as to what goes on in the integrations performed to result in the edge as seen of the SA over the three other piecewise-linear schemes used. To this end, it seems useful to look for the simplest prototype of the erosion of a step distribution. Thus, consider a four grid-interval segment of the values of q_j, with the first and the last representing the foot and the crest of the wave, respectively, having the values of 0 and 1. For convenience, let the four intervals be denoted by subscripts 1 through 4, and the unit of length be chosen equal to the grid interval. Let the second value, q_2, increase from 0 to 0.5, and at the same time the third one, q_3, decrease from 1 to 0.5, so as to keep the distribution symmetrical. In other words, we are considering the distribution

$$q_1 = 0, \quad q_2 = a, \quad q_3 = 1 - a, \quad q_4 = 1, \tag{25}$$

with values of a going from zero to 1/2. Note that, in terms of the slope ratio r of (20), this corresponds to r_2 changing from 0 to, in the limit, infinity. We are interested in slopes one obtains as this takes place, including this time also the slopes given by the superbee

$$C(r) = \max\,[0, \min\,(1,2r), \min\,(2,r)]\ . \tag{26}$$

Note that, in this simple example, N=1 is already the final state of the slope adjustment algorithm, so that further iterations would be of no impact.

As q_2 starts increasing from 0, the minmod limiter makes the slope of zone 2 immediately only half of the maximum possible for monotonicity, prevention of the spurious amplification of

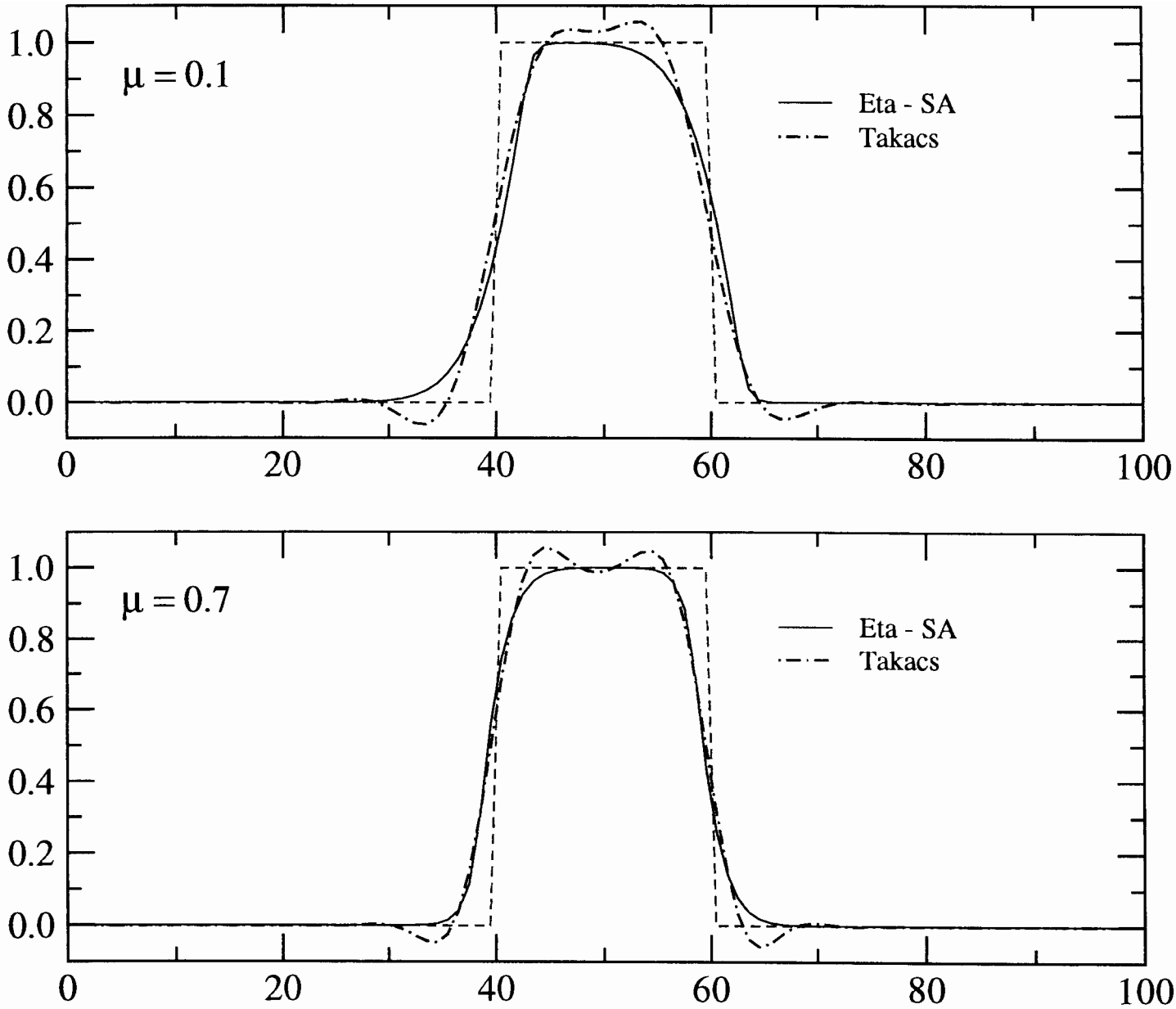

Figure 9. Same as Fig. 2, except for the Eta slope-adjustment scheme results (SA, solid line) compared against those using the Takacs (1985) third-order "minimized dissipation and dispersion errors" scheme (dot-dashed line). See text for definitions of schemes.

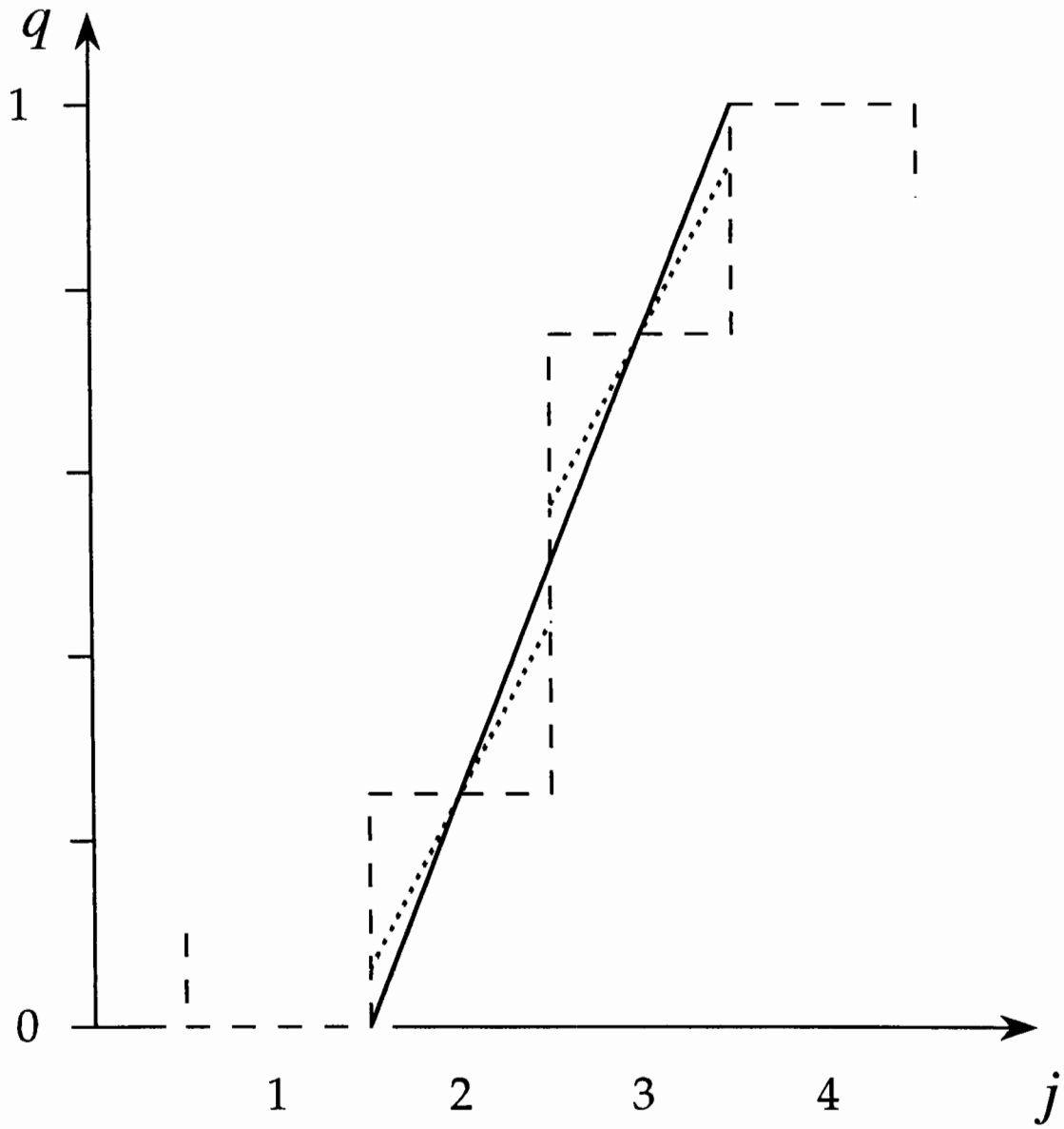

Figure 10. Distribution (25) for the case of $q_2 = 0.25$ with piecewise-constant representation (dashed line), as modified by the slope-adjustment algorithm and by the superbee limiter (solid line), and by the MC limiter (dotted lines).

extrema at this point, which is twice the difference $q_2 - q_1$. The Van Leer (harmonic average) limiter results in a greater slope, but still one smaller than twice $q_2 - q_1$. The MC and superbee limiters, as well as the SA scheme, take maximum advantage of the slope allowed for monotonicity, $2(q_2 - q_1)$. When q_2 reaches the value of 0.2, that is, for $r = 1/3$, the MC slope, still $2(q_2 - q_1)$, is the same as the centered slope, so that it is not reduced for monotonicity.

For greater values of r, the MC slope falls behind $2(q_2 - q_1)$. Thus, at $q_2 = 0.25$, for $r = 1/2$, it is $(3/2)q_2$, while the harmonic average slope is still smaller, $(4/3)q_2$, and the minmod smaller yet, q_2. At that point, the superbee and the SA slopes -- still equal to $2(q_2 - q_1)$ -- result in a straight line connecting the minimum and maximum, as shown in Fig. 10; certainly an attractive result. The MC scheme representation is in the figure shown by the dotted lines, note that its slope is 25 % smaller than it could be without violation of the monotonicity condition, with the slopes of the remaining two schemes, harmonic average and the minmod, smaller still.

At $q_2 = 1/3$, for $r = 1$, all five schemes give the same result, a slope equal to $q_2 - q_1$, resulting in a straight line but now only across the two central zones.

For still greater values of q_2, or greater values of r, the harmonic average, the MC limiter, and the superbee, in increasing intensity as ordered, depart from the straight line across the two zones, creating a sawtooth-like discontinuity at the center of the four zones. Thus, for $q_2 = 2/5$, or $r = 2$, with the SA and the minmod slopes of the two central zones remaining those of $q_3 - q_2$, the harmonic average slope in these zones is $(4/3)(q_3 - q_2)$, the MC is $(3/2)(q_3 - q_2)$, and the superbee twice $(q_3 - q_2)$, the maximum allowed for monotonicity. This situation is depicted in Fig. 11; in it the piecewise-constant representation of (25) is shown by the dashed line, the slope-adjustment and the minmod slopes are shown by the solid line, and those of the superbee by the dotted lines. The harmonic average, and the MC slopes, not shown, are between the last two, with the MC in the middle.

The steepening of the SA scheme, maximum possible while still guaranteeing no false amplification of extrema and at the same time allowing no generation of discontinuities within the piecewise-linear shape of the sign opposite to the trend of the mean values, we find appealing from the physical point of view and is the motivation for the "optimal" adjective in the title we chose for this note.

As to the point we raised of the reason for the edge of the SA over the three other piecewise-linear schemes used, this would have to be its behavior reflected in the higher steepening described for our smaller values of q_2. Specifically, the MC slope in the case depicted in Fig 10 being needlessly smaller than what the monotonicity condition would allow and what the SA slope is, would seem to be the principal reason why the SA scheme manages to perform better than the MC scheme for the square wave case as illustrated in Fig. 8.

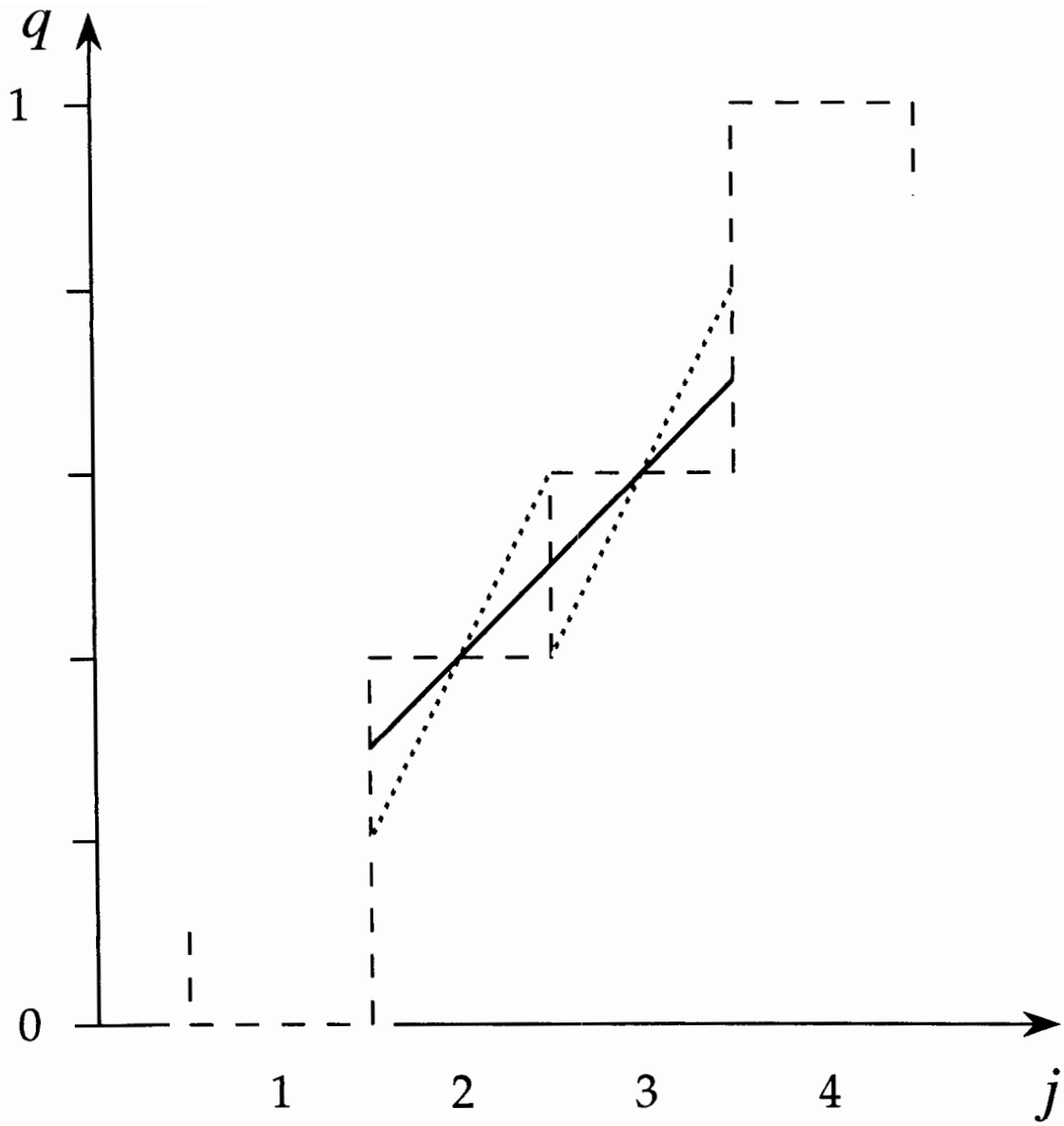

Figure 11. Distribution (25) for the case of $q_2 = 2/5$ with piecewise-constant representation (dashed line), as modified by the slope-adjustment algorithm and by the minmod limiter (solid line), and by the superbee limiter (dotted lines). Representations of the Van Leer and the MC limiters, not shown, are between the last two; with the MC in the middle, and the Van Leer two thirds of the way from superbee towards the SA/minmod solid line slope.

One remaining point regarding the performance of the SA scheme in our tests is its dependence on the number of iterations used. We have run tests of the SA scheme with our 20-zone wave case for only one, and for five iterations; results along with those for three iterations as used in the tests presented in Figs. 2 through 6 and 9 are shown in Fig. 12. A marked deterioration is seen with only one iteration compared to the result for three; but hardly any further improvement is seen to have resulted from increasing the number of iterations beyond three.

A comprehensive discussion of the issues arising for more complex advection is beyond the scope of the present note. In the Eta, the flux form of the vertical advection is retrieved from the advection form via

$$-\eta\frac{\partial q}{\partial \eta} = q\frac{\partial \eta}{\partial \eta} - \frac{\partial(\eta q)}{\partial \eta},$$

the η being the model's vertical coordinate.

Given the geometrical nature of the scheme, extension to unequal grid intervals, vertical layers in case of the Eta, is fairly straightforward. For this as well as for the extension to velocities of unprescribed signs, readers are referred to an appendix to the documentation of Black (1988). It should be kept in mind though that technical issues of these extensions are separate from the performance issues; regarding the performance with unequal grid intervals, note the recently pointed out "inherent difficulty in discretizing the advection equation" (Arakawa 2000). With a feature being advected from a region of higher resolution to a region of lower resolution, even with no phase speed error, objectives of maintaining the amplitude, conservation, and positive definiteness appear inherently at odds among themselves; a topic that has received little attention so far.

6. Concluding remarks

One objective of our note was to increase the awareness in the atmospheric modeling community of the availability of relatively simple, highly accurate, piecewise-linear schemes. By design and within the framework of the classical linear advection problem, they are stable within the CFL condition, conservative, and are easily defined so as to maintain monotonicity and guarantee no false amplification of extrema during the advection process. Note that the exact calculation of the transport step of the piecewise-linear schemes is an inherent advantage they have over their counterparts within the related family of the flux-limited schemes. The specific

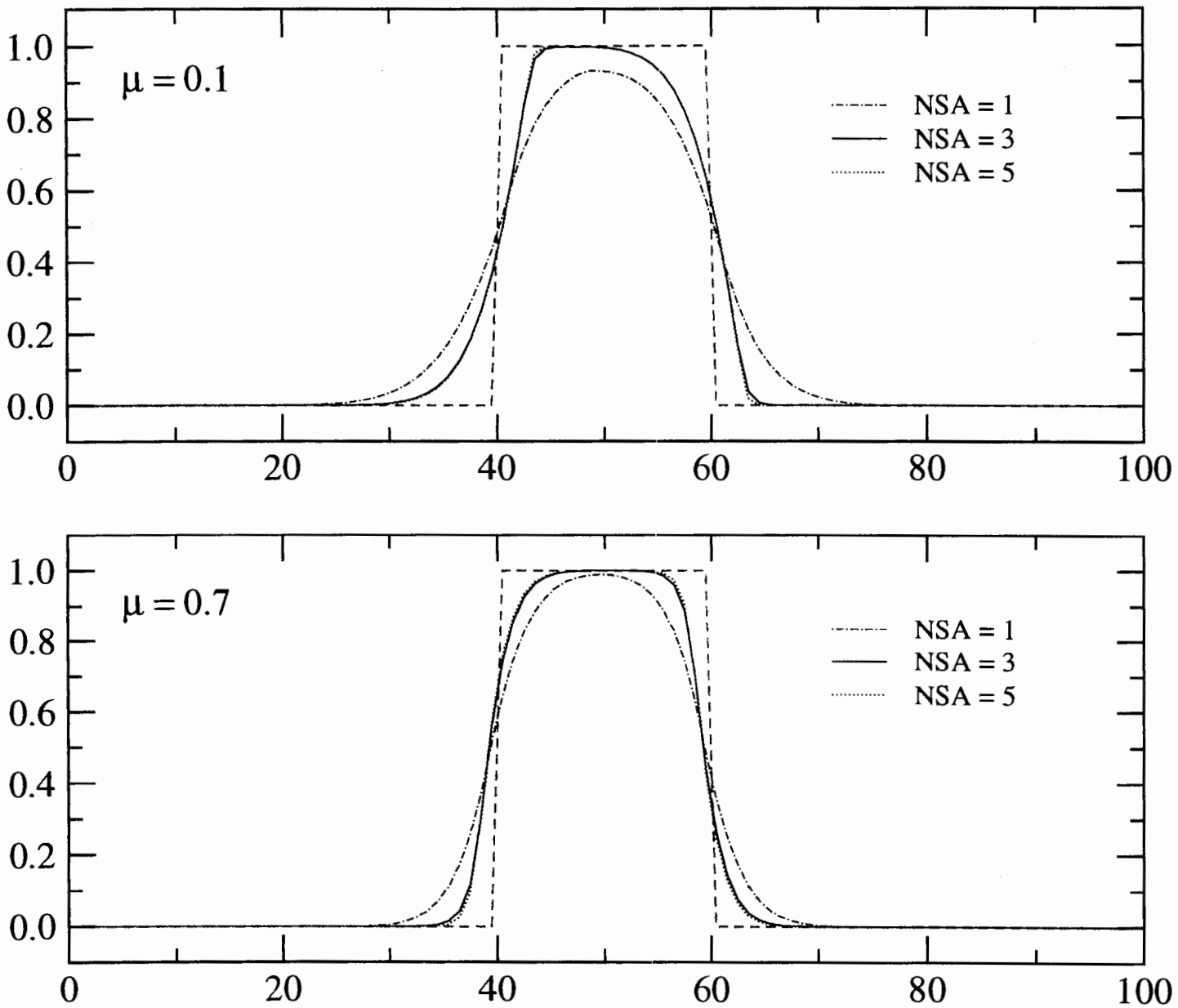

Figure 12. Same as Fig. 2, except for showing results obtained using the Eta slope-adjustment scheme run with one (dot-dashed line), three (solid line) and five (dotted line) iterations of the slope adjustment algorithm.

piecewise-linear scheme we have tested, with an iterative adjustment of slopes within the grid zones, is "strictly monotonic" – a term we suggest be added to the rich collection of terms in use already to denote that not only will there be no generation of new extrema within the monotonic segments of the original distribution, but also that the intermediate piecewise-linear representation will remain monotonic.

The slope adjustment algorithm we use takes full advantage of the steepening permitted under the condition of no spurious amplification of extrema, as long as this does not violate the stated strict monotonicity requirement. This seems to us appealing from the physical point of view. A simple example of the preceding section provided an illustration of this that we see as desirable behavior not being shared by any of the four prominent slope-limiter algorithms we have looked into. The experimental results presented are consistent with this favorable view of the algorithm used. We find particularly worthwhile noting the greater steepness obtained by the SA scheme than either that of the Van Leer or the MC algorithm scheme in the results shown in Figs. 5 and 6, respectively, in spite of the SA slopes being smaller than those of the last two in the situation shown in Fig. 11. Recall that the steepness greater than that of the SA scheme in the situation illustrated in that figure requires a violation of our strict monotonicity requirement. A problem we see with exceeding this strict monotonicity steepness threshold is that no principle appears to have been developed as to how much of a greater steepness might be permissible, if any, while still avoiding the "overcompressive" character of the solution, referred to earlier.

The issue of the economy of the scheme in present computing environments clearly depends on the machine architecture at hand. If used for the vertical advection only and the code is a massively-parallel processor (MPP) code, the cost should not be a concern if the domain decomposition is done in horizontal as normally done, since no communication is then involved. When used in horizontal, the communication overhead due to the non-local nature of the scheme will depend on the number of iterations performed, but can obviously be eliminated if desired by doing additional iterations inside the subdomains only.

We see it as desirable to test the use of the scheme for the vertical advection of all variables in the Eta given that we are aware of problems, in rare but possible situations (Rancic, personal communication), that arise from the use of the centered schemes for the vertical advection of the dynamic variables in the Eta presently.

Acknowledgments

This study has come to life as a result of cooperation between the NCEP Environmental Modeling Center, Camp Springs, MD, and the Abdus Salam International Centre for Theoretical Physics (ICTP), Miramare, Trieste, Italy; established by the directors of the the two institutions, Stephen Lord, and

Miguel Virasoro, respectively. Kind support of Filippo Giorgi, head of the ICTP's Physics of Weather and Climate Group, is much appreciated. Once again, Tom Black has given useful suggestions toward improving the manuscript. Shrinivas Moorthi and Hendrik Tolman, as EMC's internal reviewers, identified several points of the original text which needed to be clarified; Hendrik Tolman also suggested our performing the test of present Fig. 7.

The original version of the manuscript was submitted to the Quarterly Journal of the Royal Meteorological Society, which has resulted in extended discussion with Dale Durran, one of the two reviewers. This discussion led to improvements in the text of Section 4, to our discovering a boundary condition code error, leading to a slight revision of Figs. 2 through 6, 9 and 12 compared to their original versions; and finally, to our performing additional experiments with the sine plus a pulse distribution, and a "narrow" square wave case, shown in Figs, 7 and 8, respectively. These efforts however made us miss the Quarterly Journal revision deadline. Given the slowness of the peer-reviewed journals revision process and the interest we are aware of in the technique and the results presented, we find making this note available to readers in form of an NCEP Office Note desirable. The very beneficial impact of Dale Durran's input as described above is gratefully acknowledged.

References

Arakawa, A., 2000: Future development of general circulation models. *General Circulation Model Development: Past, Present and Future.* International Geophysics Series, Vol. 70, D. A. Randall, Ed., Academic Press, 721-780.

Black, T. L., 1988: The step-mountain eta coordinate regional model: A documentation. NOAA/NWS National Meteorological Center, 54 pp. [Available from NOAA Environmental Modeling Center, Room 207, 5200 Auth Road, Camp Springs, MD 20746.]

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

Carpenter, R. L., Jr., K. K. Droegemeier, P. W. Woodward, and C. E. Hane, 1990: Application of the piecewise parabolic method (PPM) to meteorological modeling. *Mon. Wea. Rev.*, **118**, 586-612.

Durran, D. D., 1999: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics.* Springer, 466 pp.

Gallus, W. A., Jr., and M. Rancic, 1996: A non-hydrostatic version of the NMC's regional Eta model. *Quart. J. Roy. Meteor. Soc.*, **122**, 499-513.

Janjic, Z. I., 1997: Advection scheme for passive substances in the NCEP Eta Model. *Res. Activities Atmos. Oceanic Modelling*, Rep. 25, WMO, Geneva, 3.14.

LeVeque, R. J., 1996: High-resolution conservative algorithms for advection in incompressible flow. *SIAM J. Numer. Anal.*, **33**, 627-665.

Mesinger, F., 1988: False precipitation enhancement due to the vertical moisture advection scheme. *Res. Activities Atmos. Oceanic Modelling*, Rep. 11, WMO, Geneva, 3.22-3.23.

Mesinger, F., and Z. I. Janjic, 1990: Numerical methods for the primitive equations (space). *10 Years of Medium-Range Weather Forecasting*, Seminar 1989, Vol. 1, ECMWF, Shinfield Park, Reading, U.K., 205-251.

Rancic, M., 1992: Semi-Lagrangian piecewise biparabolic scheme for two-dimensional horizontal advection of a passive scalar. *Mon. Wea. Rev.*, **120**, 1394-1406.

Rasch, P. J., and D. L. Williamson, 1990: Computational aspects of moisture transport in global models of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **116**, 1071-1090.

Roe, P. L., 1985: Some contributions to the modelling of discontinuous flows. *Lectures in Appl. Math.*, **22**, Part 2, 163-193.

Smolarkiewicz, P. K., 1983: A simple positive definite advection scheme with small implicit diffusion. *Mon. Wea. Rev.*, **111**, 479-486.

Sweby, P. K., 1985: High resolution TVD schemes using flux limiters. *Lectures in Appl. Math.*, **22**, Part 2, 289-309.

Takacs, L. L., 1985: A two-step scheme for the advection equation with minimized dissipation and dispersion errors. *Mon. Wea. Rev.*, **113**, 1050-1065.

Thuburn, J., 1993: Use of a flux-limited scheme for vertical advection in a GCM. *Quart. J. Roy. Meteor. Soc.*, **119**, 469-487.

Van Leer, B., 1974: Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second order scheme. *J. Comput. Phys.*, **14**, 361-370.

Van Leer, B., 1977: Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection. *J. Comput. Phys.*, **23**, 276-299.

Wurtele, M. G., 1961: On the problem of truncation error. *Tellus* , **13**, 379-391.