Advection of Coupled Hydrometeor Quantities in Bulk Cloud Microphysics Schemes

HUGH MORRISON
National Center for Atmospheric Research, Boulder, Colorado

ANDERS A. JENSEN AND JERRY Y. HARRINGTON
Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

JASON A. MILBRANDT
Atmospheric Numerical Weather Prediction Research, Environment Canada, Dorval, Canada

(Manuscript received 17 October 2015, in final form 29 February 2016)

ABSTRACT

This paper discusses the advection of coupled hydrometeor quantities by air motion in atmospheric models. It is shown that any bulk property derived from a set of advected microphysical variables must meet certain conditions in order to be preserved during transport using linear or semilinear advection schemes when the property is initially uniform, with implications for physical consistency of the property. A new, efficient flux-based method for calculating hydrometeor advection, similar to vector transport applied previously in aerosol modeling, is also presented. In this method, called scaled flux vector transport (SFVT), lead scalars (the mass mixing ratios) are advected using the host model’s unmodified advection scheme and secondary scalars (e.g., number mixing ratios) are advected by appropriately scaling the lead scalar fluxes. By design, SFVT retains linear relationships between the advected scalars. Analytic tests reveal that mean errors using SFVT are similar to those incurred using the traditional approach of separately advecting each variable. SFVT is applied to the multimoment predicted particle properties bulk microphysics scheme in idealized two-dimensional squall-line simulations using the Weather Research and Forecasting Model. The computational cost in total wall clock run time is reduced by 10%–15% while producing solutions similar to the traditional approach. Thus, SFVT can reduce the overall cost of using multimoment bulk microphysics schemes, making them competitive with simpler schemes having fewer prognostic variables.

1. Introduction

The parameterization of cloud and precipitation microphysics is a key component of atmospheric modeling. It directly impacts distributions of latent heating and cooling and condensate loading in the atmosphere, affecting buoyancy and hence dynamics. It is also critical for cloudy radiative transfer, cloud chemistry, and cloud–aerosol interaction in models. Given this importance, continuing effort has been devoted to improving microphysics parameterizations.

While bulk microphysics schemes are efficient compared to bin and Lagrangian particle schemes, extensive work over the past few decades has added considerable detail to them. This effort has included adding more prognostic variables to predict multiple moments of the particle size distribution (PSD) (e.g., Koenig and Murray 1976; Ziegler 1985; Ferrier 1994; Seifert and Beheng 2001; Morrison et al. 2005; Milbrandt and Yau 2005; Phillips et al. 2007; Lim and Hong 2010). Recent work has also added prognostic variables to improve evolution of ice particle properties (e.g., Connolly et al. 2006; Morrison and Grabowski 2008; Mansell et al. 2010; Milbrandt and Morrison 2013; Harrington et al. 2013a,b;
Morrison and Milbrandt 2015). With the trend of including two or more prognostic variables for each hydrometeor category in bulk schemes, an important question arises: How do numerical errors incurred during the physical transport of these coupled quantities affect schemes, whose representation of the PSD and process rates depend on certain relationships between the advected variables?

Previous studies have discussed two basic constraints for the choice of advected variables in cloud microphysics schemes. First, it is critical that advected microphysical variables be extensive quantities (Mansell et al. 2010; Milbrandt and Morrison 2013; Morrison and Milbrandt 2015). Directly advecting intensive quantities can generate large errors from numerical diffusion and dispersion. Thus, intensive microphysical quantities must be formulated in such a way that they can be expressed as functions of extensive quantities. A simple example is the bulk particle density, an intensive quantity, which can be expressed as the ratio of two extensive quantities: mass mixing ratio and volume mixing ratio (Connolly et al. 2006; Mansell et al. 2010; Milbrandt and Morrison 2013).

A second important constraint is consistency between the prognostic variables in a microphysics scheme and the advection algorithm in the host model when there are changes in fluid density [see Mansell (2010) for a detailed discussion]. Many models assume continuity equations for quantities that are parcel invariant, that is, quantities that are unaffected by changes in the parcel mass. For example, cloud water mass mixing ratio is the mass of cloud water per unit mass of air (kg kg$^{-1}$) and hence parcel invariant, while droplet number concentration (m$^{-3}$) is noninvariant because it is affected by parcel compression. As pointed out by Mansell (2010), inadvertent errors may arise when noninvariant microphysical quantities are advected by schemes designed to handle invariant quantities (or, less commonly, vice versa).

Beyond these two basic constraints, errors in microphysical quantities derived from two or more prognostic, advected variables in multimoment/multivariable microphysics schemes can still occur. These are quantities, for example, that describes the shape and slope of the PSD and mean particle size and fall speed in two- and three-moment schemes. Errors in these quantities in turn affect the calculation of microphysical process rates. As discussed in section 2, these errors depend upon the relationship between the derived quantities and advected variables, how the advected variables are initially distributed (specifically, if they are linearly or nonlinearly correlated), and the type of advection scheme employed.

This problem has been studied previously for aerosol microphysical models that include several advected aerosol variables (Wright et al. 2000; Wright 2007; McGraw 2007). These papers have discussed how advection can lead to unphysical features such as a negative PSD variance (Wright 2007; McGraw 2007). A few techniques have been devised to address this problem. These include vector transport (VT), where several quantities are advected together with a lead moment (Wright 2007), as well as a nonnegative least squares method in which new moment sets are resolved into sums of previously validated sets using nonnegative least squares to eliminate invalid moment sets (McGraw 2007).

Only a few studies have addressed this issue for cloud microphysics parameterizations. Sulia et al. (2013) discussed how errors in microphysical quantities derived from the ratio of prognostic advected variables are small because diffusion and dispersion errors in the two advected variables tend to cancel by taking their ratio. Ovtchinnikov and Easter (2009) showed that there can be up to ~10% error for the sum of separately advected quantities, relevant to bin microphysics, and proposed some simple corrections to reduce this error. Stevens et al. (1996) and Grabowski and Morrison (2008) discussed errors in the supersaturation derived from separate advection of water vapor mixing ratio and temperature. However, no study has looked at such errors in a more general way for multimoment/multivariable bulk schemes. The purpose of this study, therefore, is to address this gap. Specifically, we discuss functional forms of microphysical quantities derived from the set of advected variables that are preserved under certain conditions, as required for physical consistency. We also propose and test a new method to advect coupled scalar quantities similar to VT, which provides similar solutions for a reduced computational cost compared to the traditional method of independently advecting each microphysical scalar.

2. Criteria for the preservation of derived microphysical quantities

Nearly all schemes used in Eulerian atmospheric models to advect scalars apply techniques to ensure
positive definiteness. This is necessary to conserve total water mass while simultaneously avoiding negative, unphysical values of cloud quantities. From Godunov’s theorem (Godunov 1959), the only linear scheme that retains positive definiteness is first-order upwind advection, but it is far too diffusive for most applications. Higher-order schemes reduce numerical diffusion and improve accuracy, but can generate spurious oscillations and may not retain positive definiteness. Models using these schemes typically either adjust negative values of cloud quantities by “hole filling,” that is, setting negative values to zero and adjusting the field elsewhere to conserve total scalar amount, or they adjust the advective fluxes to avoid generating negative values (e.g., Zalesak 1979; Thuburn 1996; Skamarock 2006). Both of these methods introduce nonlinearity in schemes, and can increase numerical diffusion (moreover, hole filling generally does not preserve linear correlations, whereas flux limiters generally do). Atmospheric models also use higher-order, nonlinear, essentially nonoscillatory methods to limit negative values (e.g., Jiang and Shu 1996; Schroeder et al. 2006). Essentially nonoscillatory methods are nearly monotonic but do not guarantee positive definiteness; however, they can be combined with hole-filling or flux-limiting approaches to strictly ensure positive definiteness.

The transport of hydrometeors depends on air motion as well as the hydrometeor characteristics determining fall speed. However, in most models advection by air motion is calculated separately from sedimentation; in this paper we are concerned only with the former. A fundamental aspect of scalar transport following air motion in the real atmosphere is that the transport velocity is independent of the scalar, with the exception of Kolmogorov-scale mixing, which depends on diffusivity coefficients that can vary with the particular scalar being mixed (Thuburn and McIntyre 1997). This implies a certain linearity property, so that the net effect of transport can be understood as a nearly linear mathematical operator. Thus, it is desirable that advection schemes retain this linear property, while also minimizing diffusion and dispersion errors and retaining positive definiteness for water mass.

Advection schemes used in models based on the flux form of the advection equation calculate the extensive scalar quantity to be advected at the interface between grid boxes, $S_f$, with the following form:

$$S_f = \Psi(S),$$

where $S$ is the scalar at the gridbox center and the operator $\Psi(S) = B$ for a uniform field $S = B$ (Thuburn and McIntyre 1997). For linear schemes, $\Psi$ is linear and, hence,

$$AS_f + B = \Psi(AS + B).$$

Thus, by definition these schemes retain linear relationships between advected scalars, consistent with the linear property of scalar advection. For nonlinear schemes, $\Psi$ is nonlinear, however many nonlinear schemes still satisfy (2). These schemes have been referred to as “semilinear” (Lin and Rood 1996; Thuburn and McIntyre 1997; Lauritzen and Thuburn 2012), and they also preserve initial linear relationships. The nonlinear advection schemes used in this study are all semilinear. Note that semilinear and linear schemes generally do not preserve initial relationships that are nonlinear (Thuburn and McIntyre 1997; Lauritzen and Thuburn 2012).

Of particular interest to cloud microphysics modeling, the linear property of scalar advection can also be understood in terms of the evolution of intensive microphysical quantities that depend on two or more advected extensive variables. These intensive quantities include, for example, the mean particle size that depends on $(Q/N)^{1/3}$ (for spherical particles) in two-moment microphysics schemes, where $Q$ is the mass mixing ratio and $N$ is the number mixing ratio, both of which are advected. If $Q$ and $N$ are initially linearly correlated within a finite hydrometeor region and zero elsewhere, this implies $Q = AN + B$ with $B = 0$ and $A$ equal to a constant [this situation is referred to hereafter as linear cloud advection (LCA)]. Linear correlation of $Q$ and $N$ also implies that mean particle size is initially uniform across the hydrometeor region. Since linear relationships are preserved using linear and semilinear schemes following (2), this implies the ratio $Q/N$ and hence uniform mean particle size are preserved during LCA. This is important for physical consistency, since intensive microphysical quantities such as mean particle size that are initially uniform should not be affected by pure advective transport when the prognostic microphysical variables are all linearly related; any deviation from this behavior is physically inconsistent. Thus, the point of testing solutions for LCA is not because this is a realistic situation (derived intensive properties like mean volume diameter will vary spatially), but because it can reveal unphysical behavior. Note that if $Q$ and $N$ are not linearly related, then mean particle size must be initially nonuniform across the hydrometeor region and its monotonicity is not guaranteed even if $Q$ and $N$ are each advected using a monotonic scheme. This aspect is discussed further in sections 3 and 4.
Unlike mean particle size in two-moment microphysics schemes, some important intensive microphysical quantities are not a function of the ratio of two advected variables; they can depend on three or more variables in complicated, nonlinear ways. An example from the three-moment bulk microphysics scheme of Milbrandt and Yau (2005) is the shape parameter of the gamma PSD, \( \mu \). This is a function of the three advected extensive variables \( (Q, N, Z) \), approximated by

\[
\mu = a_2 G^2 + a_1 G + a_0,
\]

where \( G \) is given by

\[
G = \frac{NZ}{Q^2}.
\]

The parameter \( Z \) is the sixth moment of the PSD (proportional to radar reflectivity for spherical particles) and \( a_0, a_1, \) and \( a_2 \) are parameters for a piecewise polynomial fit (Table 1). Here \( \mu \) is a key quantity because it affects the PSD shape and hence calculation of the microphysical process rates. Without numerical testing, it is not obvious if initially uniform quantities with complicated functional dependencies on the advected variables, such as \( \mu \), are preserved during LCA. To address this question, general analytic criteria for the preservation of intensive microphysical quantities during LCA are derived next.

An intensive quantity \( F(G_s) \) is expressed as a function of some set \( G_s(S_y) \), where \( S_y = \{S_1, S_2, S_3, \ldots \} \) is the set of advected variables. Writing each \( G \in G_s \) as the ratio of two generalized power series of \( S_y \) gives

\[
G = \frac{\sum_{j=1}^{J} (\alpha_j S_1^{\beta_{1,j}} S_2^{\beta_{2,j}} S_3^{\beta_{3,j}} \ldots)}{\sum_{k=1}^{K} (\gamma_k S_1^{\phi_{1,k}} S_2^{\phi_{2,k}} S_3^{\phi_{3,k}} \ldots)},
\]

where \( J \) and \( K \) are the number of terms in the power series and \( \alpha, \beta, \gamma, \) and \( \phi \) are constants. This particular functional form for \( G \) is chosen because it is required in order for intensive quantities to be preserved during LCA, as explained below, and because many intensive microphysical quantities are functions of power laws, such as the mean particle size and \( \mu \) described above.

For LCA using linear or semilinear schemes, all advected microphysical variables are related by

\[
S_1 = A_2 S_2 = A_3 S_3 = A_4 S_4 = \ldots,
\]

where each \( A \) is a constant. Thus, in this situation \( G \) can be expressed in terms of \( S_1 \) alone by combining (5) and (6) to give

\[
G = \frac{\sum_{j=1}^{J} (\alpha_j A_2^{-\beta_{1,j}} A_3^{-\beta_{2,j}} A_4^{-\beta_{3,j}} \ldots) S_1^{\beta_{1,j} + \beta_{2,j} + \beta_{3,j} + \ldots}}{\sum_{k=1}^{K} (\gamma_k A_2^{-\phi_{1,k}} A_3^{-\phi_{2,k}} A_4^{-\phi_{3,k}} \ldots) S_1^{\phi_{1,k} + \phi_{2,k} + \phi_{3,k} + \ldots}}.
\]

In order for each \( G \) and hence \( F(G_s) \) to be preserved during LCA, \( S_1 \) must factor out of (7). If it does not, then \( G \) depends on \( S_1 \) and, therefore, suffers from diffusion and/or dispersion errors inherent in advecting \( S_1 \) itself. Therefore, preservation of an initially uniform \( F(G_s) \) during LCA requires each \( G \) to be of the functional form given by (5), and

**Table 1.** Values of parameters \( a_0, a_1, \) and \( a_2 \) from the piecewise polynomial approximation for the PSD shape parameter \( \mu \) in the three-moment bulk microphysics scheme of Milbrandt and Yau (2005). For \( G \geq 20 \) and \( G < 1.230 \) \( \mu \) is set to 0 and 40, respectively.

<table>
<thead>
<tr>
<th>( a_2 )</th>
<th>( a_1 )</th>
<th>( a_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.31 ( \leq G &lt; 20 )</td>
<td>( 3.3638 \times 10^{-3} )</td>
<td>( -1.7152 \times 10^{-1} )</td>
</tr>
<tr>
<td>7.123 ( \leq G &lt; 13.31 )</td>
<td>( 1.5900 \times 10^{-2} )</td>
<td>( -4.8202 \times 10^{-1} )</td>
</tr>
<tr>
<td>4.2 ( \leq G &lt; 7.123 )</td>
<td>( 1.0730 \times 10^{-1} )</td>
<td>( -1.7481 )</td>
</tr>
<tr>
<td>2.946 ( \leq G &lt; 4.2 )</td>
<td>( 5.9070 \times 10^{-1} )</td>
<td>( -5.7918 )</td>
</tr>
<tr>
<td>1.793 ( \leq G &lt; 2.946 )</td>
<td>( 4.3966 )</td>
<td>( -2.6659 \times 10^{1} )</td>
</tr>
<tr>
<td>1.405 ( \leq G &lt; 1.793 )</td>
<td>( 4.7552 \times 10^{1} )</td>
<td>( -1.7958 \times 10^{2} )</td>
</tr>
<tr>
<td>1.230 ( \leq G &lt; 1.405 )</td>
<td>( 3.0889 \times 10^{2} )</td>
<td>( -9.0854 \times 10^{2} )</td>
</tr>
</tbody>
</table>

**Fig. 1.** An algorithmic description of the advection of mass mixing ratio \( Q \) (lead scalar) and nonmass microphysical variables \( S \) (secondary scalars) using SFVT as implemented into WRF with a three-step Runge-Kutta (RK) solver.
For each \( j \) from 1 to \( J \) and each \( k \) from 1 to \( K \), the diffusion and/or dispersion errors exactly cancel so that the quantities are preserved. As long as each \( G \in G_x \) is of the form in (5) and meets the criteria in (8), then \( F \) can be any real function of the set \( G_x \) and remain invariant for LCA.

The simplest nontrivial (nonconstant) function \( G \) consistent with (5)–(8) is the ratio of two advected quantities. This includes the mean particle size as a function of \( G^{1/3} \), where \( G = Q/N \), as discussed above. Other examples relevant to microphysics include the bulk density equal to the ratio of \( Q \) and the particle volume mixing ratio (Connolly et al. 2006; Mansell et al. 2010; Milbrandt and Morrison 2013; Morrison and Milbrandt 2015), crystal axis ratio equal to the ratio of the \( c \)- and \( a \)-axis length mixing ratios (Harrington et al. 2013a,b; Sulia et al. 2013), and the rime mass fraction equal to \( Q_{\text{rim}}/(Q_{\text{rim}} + Q_{\text{dep}}) \), where \( Q_{\text{rim}} \) and \( Q_{\text{dep}} \) are the ice mass mixing ratios grown by riming and vapor deposition, respectively (Morrison and Grabowski 2008).

### Table 2: Overview of the configurations used in the analytic 1D tests

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRF-PD</td>
<td>TRAD using the standard WRF positive-definite scheme</td>
</tr>
<tr>
<td>WRF-PD-SFVT</td>
<td>SFVT with fifth-order weighting using the standard WRF positive-definite scheme</td>
</tr>
<tr>
<td>WRF-MONO</td>
<td>TRAD using the standard WRF monotonic scheme</td>
</tr>
<tr>
<td>WRF-MONO-SFVT</td>
<td>SFVT with fifth-order weighting using the standard WRF monotonic scheme</td>
</tr>
<tr>
<td>WENO</td>
<td>TRAD using the WENO scheme</td>
</tr>
<tr>
<td>WENO-SFVT</td>
<td>SFVT with fifth-order weighting using the WENO scheme</td>
</tr>
<tr>
<td>WRF-MONO5</td>
<td>Advection of ( S_1 ) using the standard WRF monotonic scheme; advection of ( S_2 ) using the standard linear S08 scheme</td>
</tr>
<tr>
<td>WRF-MONO1</td>
<td>Advection of ( S_1 ) using the standard WRF monotonic scheme; advection of ( S_2 ) using the first-order upwind method</td>
</tr>
</tbody>
</table>

### Table 3: Error statistics for the 1D analytic advection tests

<table>
<thead>
<tr>
<th>Case</th>
<th>WRF-PD</th>
<th>WRF-PD-SFVT</th>
<th>WRF-MONO</th>
<th>WRF-MONO-SFVT</th>
<th>WENO</th>
<th>WENO-SFVT</th>
<th>WRF-MONO5</th>
<th>WRF-MONO1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRI1</td>
<td>1.94</td>
<td>1.21</td>
<td>0.81</td>
<td>1.42</td>
<td>1.19</td>
<td>1.23</td>
<td>1.89</td>
<td>4.04</td>
</tr>
<tr>
<td>TRI2</td>
<td>1.97</td>
<td>1.38</td>
<td>2.32</td>
<td>1.28</td>
<td>1.55</td>
<td>1.58</td>
<td>1.62</td>
<td>10.33</td>
</tr>
<tr>
<td>STAIR</td>
<td>6.82</td>
<td>4.74</td>
<td>3.86</td>
<td>3.57</td>
<td>4.31</td>
<td>4.55</td>
<td>6.48</td>
<td>14.11</td>
</tr>
</tbody>
</table>
advected variables must be recast so that each linearly related, as well as during LCA, the set of three or more advected variables and they are not all initially uniform property MPDATA (not shown). If one wishes to preserve an velocity using either first-order upwind advection or confirmed by 1D numerical tests with constant flow $Z_{\text{less}}^m$, because, in contrast to LCA, (6) no longer holds in two of the variables are not linearly related. This is factor out of (5) for each $S$. Thus, an initially uniform $\mu$, while invariant during LCA, is not preserved unless $Z, N,$ and $Q$ are all linearly related, which was confirmed by 1D numerical tests with constant flow velocity, using a range of Courant numbers and either linear first-order upwind advection or the non-linear multidimensional positive-definite advection transport algorithm (MPDATA; Smolarkiewicz 1984; Smolarkiewicz and Margolin 1998). If an initially uniform derived quantity is invariant during LCA using linear or semilinear advection schemes following (5)–(8), it is not preserved if it depends on three or more advected variables and any two of the variables are not linearly related. This is because, in contrast to LCA, (6) no longer holds in this situation and, therefore, $S_1, S_2, S_3, \ldots$ cannot all factor out of (5) for each $G$. Thus, an initially uniform $\mu$, while invariant during LCA, is not preserved unless $Z, N,$ and $Q$ are all linearly related, which was confirmed by 1D numerical tests with constant flow velocity, using either first-order upwind advection or MPDATA (not shown). If one wishes to preserve an initially uniform property $F(G)$ when it depends on three or more advected variables and they are not all linearly related, as well as during LCA, the set of advected variables must be recast so that each $G \in G_x$ can be expressed as the ratio of two advected quantities. For example, if $F = G = S_1/\sqrt{S_2S_3}$, one can set $S_4 = \sqrt{S_2S_3}$ and advect $S_1$ and $S_4$ to preserve an initially uniform $F$ using linear or semilinear schemes even when $S_1, S_2,$ and $S_3$ are not all linearly correlated. The term $S_2$ (or $S_3$) can also be advected and used together with $S_4$ to derive $S_3$ (or $S_2$), maintaining consistency among the variables.

### 3. Alternative approaches for advection of coupled scalar quantities

#### a. Discussion of previous methods

The linear property of scalar advection implies the preservation of certain functional forms for microphysical quantities derived from advected scalars, as discussed in section 2. It also suggests that only a single scalar need be advected, and the advection of other scalars related to it can be calculated by a linear scaling of the advection operator. This idea has served as the basis for an approach used to advect scalar quantities representing various moments of the aerosol size distribution, the so-called VT method (Wright et al. 2000; Wright 2007; McGraw 2007). For VT, first a normalizing scalar is chosen. All other prognostic scalars related to the lead scalar are normalized by the lead scalar to yield the intensive properties of interest. Then, the lead scalar is advected using the full advection scheme. The fluxes and initial values of the lead scalar are used to determine fractional contributions $f$ from each grid box to the updated lead scalar of box $i$. Assuming operator splitting in three dimensions so that the calculation can be expressed in a single dimension, the derived intensive properties $\Pi$ are then updated as (Wright 2007)

$$\Pi(t + \Delta t) = f_{i-1} \Pi_{i-1}(t) + f_i \Pi_i(t) + f_{i+1} \Pi_{i+1}(t),$$

where the grid stencil (the grid region used to calculate fluxes at any single grid point, typically depending on the order of advection) is assumed to cover $i - 1, i,$ and $i + 1$ (note the grid stencil will depend on the specific advection scheme being employed). Finally, the updated

<table>
<thead>
<tr>
<th>Case</th>
<th>WRF-PD</th>
<th>WRF-PD-SFVT</th>
<th>WRF-MONO</th>
<th>WRF-MONO-SFVT</th>
<th>WENO</th>
<th>WENO-SFVT</th>
<th>WRF-MONO5</th>
<th>WRF-MONO1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR1</td>
<td>4.62</td>
<td>6.06</td>
<td>1.84</td>
<td>4.64</td>
<td>1.76</td>
<td>3.37</td>
<td>4.59</td>
<td>4.80</td>
</tr>
<tr>
<td>TR2</td>
<td>8.16</td>
<td>9.03</td>
<td>3.48</td>
<td>3.42</td>
<td>3.43</td>
<td>3.44</td>
<td>4.32</td>
<td>11.65</td>
</tr>
<tr>
<td>STAIR</td>
<td>12.71</td>
<td>12.62</td>
<td>3.32</td>
<td>3.26</td>
<td>3.34</td>
<td>4.88</td>
<td>9.32</td>
<td>14.52</td>
</tr>
</tbody>
</table>
prognostic extensive variables are rederived from the updated intensive quantities and the lead scalar.

Vector transport preserves valid sets of derived properties since the set of properties after advection is a linear combination of the properties before advection and the coefficients $f$ are, by definition, nonnegative. This implies that the derived properties are locally monotonic within the grid stencil. Vector transport is also computationally efficient since the full flux calculations using the model's advection scheme are only required for the lead scalar. However, VT can produce errors if spatial gradients of the lead and secondary scalars differ markedly. For example, if the lead scalar is initially uniform everywhere, its flux divergence will be zero and, hence, the fractional contributions of each grid cell to the updated value of the lead scalar will be undefined. In tests by McGraw (2007), VT led to a substantial loss of accuracy compared to applying the full advection calculation to each scalar independently. Vector transport also requires explicit calculation of the flow-dependent contributions to the scalar flux divergence from each grid cell, which requires reverse engineering of the host model's advection scheme. This is a considerable task that is specific to the particular advection scheme used by the model, and, hence, is cumbersome to implement. This is particularly true for higher-order advection schemes, in which case the contribution to the flux divergence in a grid cell may depend in complicated ways upon a fairly large number of neighboring grid cells.

To address these points with VT, McGraw (2007) proposed a nonnegative least squares method to preserve valid sets of derived intensive properties. In this approach, at each time step all extensive scalars are updated using the model's standard advection scheme applied independently to each scalar. Results are then combined to form an intermediary set of properties. Inconsistencies in this set are then resolved using coefficients from a nonnegative least squares fit to all relevant combinations of the quantities derived from the set of advected extensive scalars. The nonnegative least squares method is analogous to VT if the coefficients are derived only from a single scalar. This approach improves accuracy and reduces diffusivity compared to VT (McGraw 2007), but loses the benefit of computational efficiency.

b. Scaled flux vector transport

An alternative approach is proposed here, which we refer to as scaled flux vector transport (SFVT). It involves scaling of the advective fluxes of the lead scalar quantities, instead of calculating gridcell contributions to the updated value of the lead scalar as in VT. This generalizes the idea of applying the same positivity (or monotonic) flux correction to

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRF-S08</td>
<td>TRAD using the standard linear S08 scheme applied to all scalars</td>
</tr>
<tr>
<td>WRF-PD</td>
<td>TRAD using the WRF positive-definite scheme applied to all scalars</td>
</tr>
<tr>
<td>WRF-PD-5-SFVT</td>
<td>WRF positive-definite scheme applied to cloud, rain, and ice mass mixing ratios, and SFVT with fifth-order weighting in the horizontal and third-order weighting in the vertical used for all other scalars</td>
</tr>
<tr>
<td>WRF-PD-3-SFVT</td>
<td>WRF positive-definite scheme applied to cloud, rain, and ice mass mixing ratios, and SFVT with third-order weighting in the horizontal and vertical used for all other scalars</td>
</tr>
<tr>
<td>WRF-PD-5M-SFVT</td>
<td>As in WRF-PD-5, but including the simple limiter to preserve monotonicity of the lead and secondary scalar ratios</td>
</tr>
<tr>
<td>WENO</td>
<td>TRAD using the WENO scheme applied to all scalars</td>
</tr>
<tr>
<td>WENO-5-SFVT</td>
<td>WENO scheme applied to cloud, rain, and ice mass mixing ratios, and SFVT with fifth-order weighting in the horizontal and vertical used for all other scalars</td>
</tr>
<tr>
<td>WENO-3-SFVT</td>
<td>WENO scheme applied to cloud, rain, and ice mass mixing ratios, and SFVT with third-order weighting in the horizontal and vertical used for all other scalars</td>
</tr>
<tr>
<td>WSM6</td>
<td>TRAD using the WRF positive-definite scheme applied to all scalars in the WSM6 microphysics scheme</td>
</tr>
<tr>
<td>THOMPSON</td>
<td>TRAD using the WRF positive-definite scheme applied to all scalars in the Thompson microphysics scheme</td>
</tr>
</tbody>
</table>
multiple scalars in order to preserve initial linear correlations proposed by Blossey and Durran (2008).

SFVT is described as follows. Similar to VT, a “lead” scalar is chosen and advected using the model’s unmodified advection scheme, typically with flux limiters or other techniques to ensure positive definiteness and/or monotonicity. Given the importance of both conserving total water and avoiding negative values of hydrometeor mass mixing ratios, it is recommended that the hydrometeor mass mixing ratio $Q$ be used as the lead scalar for each hydrometeor category. Using $Q$ also guarantees the existence of nonzero lead scalar values for all grid points containing hydrometeors. Otherwise, if the lead scalar was zero in a region containing hydrometeors, such that its fluxes were also zero, the lead scalar fluxes would not provide sufficient information for calculating secondary scalar fluxes. The fluxes of secondary scalars at the grid interfaces are calculated from scaling the $Q$ fluxes by $W(S)/W(Q)$, where $W$ is a linear weighting function calculated at the grid interfaces and $S$ is the secondary scalar(s) related to $Q$ to be advected, such as the number mixing ratio in two-moment schemes. The choice of $W$ is described below. The time tendencies of $S$ from advection are then obtained by a standard flux divergence calculation of these scaled fluxes. An algorithmic description of SFVT implemented into the Weather Research and Forecasting (WRF) Model is shown in Fig. 1.

It is necessary that $W$ is a linear function of the scalar over some grid stencil, similar in form to a linear advective flux operator; otherwise, initial linear correlations

---

**Fig. 3.** Vertical cross sections of extensive advected prognostic variables at 4 h for the WRF-PD test: (a) rain mass mixing ratio $Q_r$, (b) rain number mixing ratio $N_r$, (c) ice mass mixing ratio $Q_i$, (d) ice number mixing ratio $N_i$, (e) rime ice mass mixing ratio $Q_{rim}$, and (f) rime volume mixing ratio $B_{rim}$. 
between scalars would not be preserved. Thus, by design SFVT with a linear \( W \) preserves initial linear correlations. In principle any linear \( W \) can be used, but for strict preservation of initial linear correlations \( W \) should cover the same grid stencil as the advective flux calculation for \( Q \). This is because higher-order advection schemes can generate nonzero fluxes several grid points away from any grid with nonzero \( Q \) (but still within the grid stencil). Thus, if \( W \) does not cover the same stencil as the \( Q \)-flux calculation, it is possible to have a nonzero flux of \( Q \) but a zero flux of \( S \) at the same location, leading to an inconsistency. However, in practice an exact match of the grid stencil for \( W \) and that for the flux calculation of \( Q \) does not seem necessary, as demonstrated in the tests described in section 4c. It is emphasized that \( W \) does not have to be of the same form as the flux calculation of \( Q \), which underpins the main advantage of this approach—its computational efficiency. Thus, a simple linear \( W \) can be used together with much more complicated, nonlinear flux calculations of \( Q \). Three different formulations for \( W \) are described in section 4a.

SFVT does not guarantee monotonicity of intensive quantities if they are initially nonuniform in a region with \( Q > 0 \), unlike VT. This is demonstrated by a simple example in 1D nondivergent flow where the field of \( Q \) is constant, while \( S \) varies. In this case, the fluxes of \( Q \) into and out of a grid box must be equal. Using SFVT, we can write the flux divergence of \( S \) as

\[
\text{div}(S) = \frac{1}{\Delta x} \left[ F_{i+1/2}(Q) \frac{W_{i+1/2}(S)}{W_{i+1/2}(Q)} - F_{i-1/2}(Q) \frac{W_{i-1/2}(S)}{W_{i-1/2}(Q)} \right],
\]  

(10)
where $F(Q)$ is the flux of $Q$ at the grid interfaces, $W$ are the weighting functions corresponding to flux calculations, and $\Delta x$ is the grid spacing. Since $Q$ is constant, we have $W(Q) = W_{i+1/2}(Q) = W_{i-1/2}(Q)$, and since $W$ is a linear weighting function analogous to a linear advective flux operator, following (2) we can write $AW(S) = W(AS)$, where $A = F(Q) / W(Q)$ is constant. Thus, the flux divergence of $S$ is \[ |W_{i+1/2}(AS) - W_{i-1/2}(AS)|/\Delta x. \] It follows that since $Q$ is constant, monotonicity of $S$ and hence $S/Q$ are preserved only if $W$ is a monotonic operator. This is generally not the case for any linear $W$ except for that analogous to a first-order upwind flux calculation. In general, monotonicity of derived properties is also not preserved in the traditional approach of advecting each scalar independently, even when limiters are applied that preserve monotonicity for individual scalars. This is a major difference between SFVT and VT; the latter guarantees monotonicity of derived quantities, but can also produce inconsistent solutions as noted above. The effects of preserving monotonicity of the ratio $S/Q$ are discussed in section 4c.

4. Results

a. Description of the advection schemes tested

SFVT was tested for 1D analytic cases and two-dimensional (2D) idealized squall-line simulations using several advection options available in WRF. All of the schemes assume dimensional operator splitting. The standard WRF configuration uses a third-order Runge–Kutta time integration scheme with second- through sixth-order accuracy options (Skamarock et al. 2008). The basic linear WRF scheme (referred to hereafter as S08), following from Wicker and Skamarock (2002), is not positive definite. If negative values are produced by advection in WRF, these values are simply reset to zero meaning that total water is not conserved. To address this problem, a positive-definite (PD) flux limiter (Skamarock 2006) was later introduced into WRF (this scheme is referred to as WRF-PD hereafter). It renormalizes fluxes to prevent negative values from occurring. More recently, this scheme has been modified to include a monotonic flux limiter (Wang et al. 2009; WRF-MONO hereafter). The flux limiters are only applied to the third step of the Runge–Kutta time integration scheme, while the linear S08 scheme is applied for the first two steps. In the following tests, third-order accuracy was used in the vertical and fifth-order in the horizontal, which is typically used in WRF, giving grid stencils of three points in the vertical and five in the horizontal for the flux calculations. Since the nonlinear flux calculations in WRF are only applied to the third step of the Runge–Kutta scheme in WRF, SFVT is also only applied during the third step (see Fig. 1). Thus, it is only the grid stencil of the third step that is relevant for SFVT, although as shown by tests in section 4c a strict match of grid stencils between the nonlinear $Q$ flux calculations and $W$ is not necessary in practice.

More recently, a fifth-order weighted essentially nonoscillatory scheme (WENO) (Jiang and Shu 1996; Shu 1999) was implemented into WRF. It achieves near-positive definiteness and monotonicity through the use of polynomial reconstruction of cell boundary values for calculating fluxes. Similar to the PD and monotonic schemes, WENO is only applied to the third Runge–Kutta step in WRF, while the linear S08 scheme is applied during the first two steps. This version of WENO has grid stencils of five points each in the horizontal and vertical for the flux calculations.

Three different linear $W(X)$ were tested with SFVT:

$$W_{i-1/2}^{(\text{First})} = \frac{1}{2}(V_{i-1/2} + |V_{i-1/2}|)X_{i-1} + \frac{1}{2}(V_{i-1/2} + |V_{i-1/2}|)X_i, \quad (11)$$

$$W_{i-1/2}^{(\text{Third})} = V_{i-1/2} \left\{ \frac{7}{12} (X_i + X_{i-1}) - \frac{1}{12} (X_{i+1} + X_{i+2}) + \frac{1}{12} \frac{|V_{i-1/2}|}{V_{i-1/2}} [(X_{i+1} + X_{i+2}) - 3(X_i + X_{i-1})] \right\}, \quad (12)$$

$$W_{i-1/2}^{(\text{Fifth})} = V_{i-1/2} \left\{ \frac{37}{60} (X_i + X_{i-1}) - \frac{2}{15} (X_{i+1} + X_{i+2}) + \frac{1}{60} (X_{i+2} + X_{i-3}) ight.$$

$$- \frac{1}{60} \frac{|V_{i-1/2}|}{V_{i-1/2}} [(X_{i+2} - X_{i-3}) - 5(X_{i+1} - X_{i+2}) + 10(X_i - X_{i-1})] \right\}, \quad (13)$$

where $i$ is the grid spatial index and $V$ is the air velocity (positive for increasing $i$). The fifth-order and third-order weightings are analogous to the respective fifth- and third-order flux calculation operators in the linear S08 advection scheme based on Taylor series expansions of the flux divergences in Wicker and Skamarock (2002); $W_{i-1/2}^{(\text{First})}$ simply follows from a first-order upwind flux calculation. For the implementation of SVFT with
WENO, the growth of computational round-off error led to noisy solutions, if left unchecked. This problem can be mitigated by the use of double precision values and operations, but this is cumbersome to employ in WRF. Instead, rounding the ratio \( W(S)/W(Q) \) to the third significant digit produced satisfactory results.

b. 1D analytic tests

The accuracy of SFVT was tested using 1D analytic cases where advective transport in constant flow is considered. Results using SFVT are compared to those using the traditional approach of independently advecting each scalar (TRAD), with analytic solutions providing a benchmark. Three cases with different initial profiles of the two advected scalars (\( S_1 \) and \( S_2 \)) were tested (Fig. 2): triangular for \( S_1 \) and top-hat for \( S_2 \) (TRI1), triangular for \( S_2 \) and top-hat for \( S_1 \) (TRI2), and a two-step stairlike profile (STAIR). Additional tests with cosine initial profiles for either \( S_1 \) or \( S_2 \) give similar overall results. For SFVT, \( S_1 \) and \( S_2 \) were taken as the lead and secondary scalars, respectively. Three advection schemes were employed (WRF-PD, WRF-MONO, and WENO), with a pair of numerical solutions using SFVT or TRAD for each combination of advection scheme and test case. A list of the configurations is provided in Table 2. The test setup includes 100 grid points, and the initial hydrometeor region covers 7 grid points (grids 25–31, with flow from small to large grid number). A range of Courant numbers was tested, modified by changing the grid spacing while holding the flow velocity and time step fixed. For simplicity a forward Euler time integration scheme is employed and the full nonlinear flux calculations or SFVT are calculated and applied every step, in contrast to the third-order Runge–Kutta scheme in WRF. Consistent with the adjustment of secondary scalars within multimoment microphysics schemes to ensure that derived microphysical quantities remain within certain bounds, at each step values of \( S_2 \) were clipped if necessary by setting \( S_2 = 0 \) if \( S_2 < 0 \) and \( S_2 = S_1 \) if \( S_2 > S_1 \). Both \( S_1 \) and \( S_2 \) were set to 0 for points with \( S_1 < 10^{-13} \).

Results are illustrated in Fig. 2 for tests using the WRF monotonic advection scheme and a Courant number of 0.1; error statistics for all tests relative to the analytic solutions are presented in Tables 3–5. Results are calculated after 50 time steps using a Courant number of 0.1 in Table 3. Tables 4 and 5 show results for Courant numbers of 0.5 and 0.02 after 10 and 250 time steps, respectively. These times were chosen for analysis since the analytic solutions cover the same grid locations for the various Courant numbers tested. Analyses at other times give similar overall results.
Since SFVT and TRAD advect $S_1$ in the same way using the chosen advection scheme, errors in its evolution are identical using either method, and, therefore, are not shown in Tables 3–5. On the other hand, advection of $S_2$ differs between SFVT and TRAD. Overall, SFVT gives similar mean errors compared to TRAD for $S_2$. Also shown are error statistics from a configuration advecting $S_1$ using WRF-MONO and $S_2$ using WRF-S08 (WRF-MONO5 in Tables 3–5). This test is similar to WRF-MONO-SFVT, except that the fifth-order linear fluxes from WRF-S08 are used directly to advect $S_2$, instead of being used as weighting functions using SFVT. Mean errors in WRF-MONO5 are considerably larger than in WRF-MONO-SFVT, indicating the benefit gained by the scaling method in SFVT as opposed to using the linear fluxes directly for advecting $S_2$. Mean error is further increased in a configuration similar to WRF-MONO5 but using first-order upwind advection directly for $S_2$ instead of fifth-order linear advection (WRF-MONO1 in Tables 3–5), indicating improved accuracy using a higher-order scheme in these tests.

Results from additional tests without any clipping applied to $S_2$ (and Courant number of 0.1) are shown in Table 6. In these tests, negative values of $S_2$ are produced using SFVT with higher-order linear weighting since it does not preserve monotonicity, while negative values are avoided using the traditional approach with a positive-definite scheme. Mean error for $S_2$ is usually (though not always) larger without clipping for a given configuration, while the general picture of similar errors using SFVT and the traditional still holds. An exception is the TRI2 initial profile of $S_1$ and $S_2$ case, for which WENO-SFVT produces larger mean error (by about an order of magnitude) compared to the other tests. This occurs because relatively large magnitude negative values of $S_2$ are produced within the first several steps, although these errors do not grow in time after the first $\sim 10$ steps. These errors are of little practical importance since WENO-SFVT produces results very similar to TRAD for the WRF squall-line tests (that apply clipping of negative values for secondary scalars, as is standard within microphysics schemes), described in section 4c.

c. 2D squall-line tests

The WRF simulations employ a setup similar to the standard idealized 2D squall-line case in WRF. Squall lines provide an ideal test case because they encompass hydrometeors with a wide variety of characteristics and growth modes, and the dynamical and microphysical evolution strongly depends on derived intensive properties such as density and fall speed (e.g., Morrison et al. 2015). WRF is a compressible, nonhydrostatic atmospheric model with subtime stepping to maintain stability for acoustic modes. Here the domain is 500 km in the
horizontal and 20 km in the vertical. The horizontal grid spacing is 1 km, with 80 vertical levels using a spacing slightly stretched from 225 to 290 m. To avoid complications in interpreting results, all explicit subgrid-scale diffusion is turned off (there is still numerical diffusion from the advection). For simplicity, radiation and surface–atmosphere exchange is neglected, and free-slip conditions are employed at the model top and bottom. Lateral boundary conditions are open. A Rayleigh damper with a time scale of 0.003 s\(^2\) is used in the top 5 km of the domain to limit spurious wave reflection. Convection is initiated by applying a warm bubble with a radius of 4 km horizontally and 1.5 km vertically, centered at a height of 1.5 km, and with a maximum perturbation potential temperature of 3 K. The initial thermodynamic environment is horizontally homogeneous and follows from the analytic sounding of Weisman and Klemp (1982). An initial vertical shear of the horizontal wind of 0.0048 s\(^2\) is applied between 0 and 2.5 km. The model is integrated for 4 h using a 5-s time step.

Unless otherwise indicated, simulations use the multimoment predicted particle properties (P3) bulk microphysics scheme (Morrison and Milbrandt 2015). The version of P3 employed here includes prognostic mass mixing ratios for cloud water \(Q_c\), rain \(Q_r\), and ice \(Q_i\), the number mixing ratio for rain \(N_r\), and number \(N_i\), rime mass \(Q_{rim}\), and rime volume \(B_{rim}\) mixing ratios for ice. We take \(Q_c\), \(Q_r\), and \(Q_i\) as the lead scalars and \(N_r\), \(N_i\), \(Q_{rim}\), and \(B_{rim}\) as the secondary scalars for the tests using SFVT.

An overview of the main WRF tests is provided in Table 7. For brevity we focus on results using the WRF-PD and WENO advection schemes. Storm evolution is broadly similar for most, but not all, of the tests. Moist convection is initiated in all simulations within the first 5 min, and precipitation reaches the surface within 20 min. Widespread stratiform precipitation develops after 2–3 h, so that the simulated storm has a leading convection-trailing stratiform squall-line structure. A comparison of the simulations focuses on extensive and derived intensive microphysical quantities, as well as important model outputs such as surface precipitation rate \(q_s\) and radar reflectivity \(Z\). Derived intensive microphysical quantities analyzed are rime mass fraction \(F_r\) and the mass-weighted mean particle density \(\rho_m\) and fall speed \(V_m\). The variable \(F_r\) is simply the ratio of the prognostic \(Q_{rim}\) and \(Q_i\), while \(\rho_m\) and \(V_m\) are functions of \(Q_c/N_r\), \(Q_{rim}/Q_r\), and \(Q_{rim}/B_{rim}\).

SFVT gives very similar results compared to the corresponding simulations using the traditional approach. This is illustrated, for example, by comparing vertical cross-sectional plots of the prognostic \(Q_r\), \(N_r\), \(Q_i\), \(N_i\), \(Q_{rim}\), and \(B_{rim}\) from WRF-PD and WRF-PD-5-SFVT (Figs. 3 and 4). (Note \(Q_c\) is not shown since the cloud water category has no secondary scalars in this version of the P3 scheme.) There is a similar correspondence.

**Fig. 7.** As in Fig. 5, but for the WRF-PD-5-SFVT test.
between WENO and WENO-5-SFVT (not shown). Thus, WRF-PD-5-SFVT and WENO-5-SFVT produce almost the same structure of the intensive properties compared to WRF-PD and WENO, respectively (cf. Figs. 5–6 and 7–8). Rime mass fraction $F_r$ is large along the leading edge of convection, where there is substantial riming, and small in the stratiform precipitation region, where ice growth is dominated by vapor deposition; $\rho_m$ is large ($>700 \text{ kg m}^{-3}$) in two distinct regions: at low and midlevels along the leading convective line and at upper levels in the anvil region; and $V_m$ is greatest at lower levels in the convective region and decreases with height and distance from the leading edge of convection. Morrison and Milbrandt (2015) and Morrison et al. (2015) provide further analysis of squall-line simulations using WRF with the P3 scheme.

Similarity between the tests using SFVT and the corresponding tests using the traditional approach is further illustrated by histograms of $Z$, surface precipitation rate $Q_i$, and $F_r$ (Fig. 9). Notably, there do not appear to be any systematic errors or biases. Differences are largest for values of $F_r$ less than about 0.2, but these do not have much impact on overall simulations. WRF-PD and WENO give rather different solutions, larger than the differences between SFVT and the corresponding tests using the traditional approach. Thus, SFVT is able to preserve differences between WENO and WRF-PD even though the full advective flux calculations are only applied to the mass mixing ratios when using SFVT. In these simulations, WRF-PD produces faster storm evolution and a larger area of hydrometeors compared to WENO.

Tests using SFVT with third-order weighting in both the horizontal and vertical and either WRF-PD or WENO (WRF-PD-3-SFVT, WENO-3-SFVT) (Figs. 10 and 11) are also similar to those using the traditional approach, although differences between WENO and WENO-3-SFVT are somewhat greater than between WENO and WENO-5-SFVT. Nonetheless, it is interesting that WRF-PD-3-SFVT and WENO-3-SFVT produce results quite close to the corresponding WRF-PD and WENO simulations despite some mismatch in grid stencil between $W_{i-1/2}^{(\text{Third})}$ and the flux calculations for the lead scalars. On the other hand, using $W_{i-1/2}^{(\text{First})}$ with either the WRF-PD or WENO schemes, while computationally very efficient, produced unsatisfactory results and noisy intensive property fields (not shown). This occurred because of the rapid growth of the computational round-off error, and because of the large mismatch between the first-order weighting grid stencil and that of the lead scalar flux calculations.

Solutions are considerably different using the linear S08 scheme to advect all scalar quantities independently (Fig. 12). Besides not conserving total water mass, there is convection initiated well away from the surface gust.
front in the rearward part of the storm leading to a disorganized structure, and the leading edge of the storm is ~50 km ahead of the WRF-PD simulation. Given the efficiency of the linear S08 scheme, it is reasonable to wonder if WRF-PD or WENO could be used to advect the lead scalars and the linear flux calculations in S08 to advect the secondary scalars directly. However, this approach leads to divergence of solutions compared to SFVT and noisy fields for the derived microphysical properties, especially using WENO. It also violates the linear property of scalar advection, and, hence, does not preserve initial linear correlations, and as discussed in section 4b leads to a decrease in accuracy compared to SFVT. Thus, it is important to properly weight the lead scalar fluxes for calculating advection of secondary scalars in SFVT.

As mentioned earlier, SFVT does not guarantee monotonicity of ratios of secondary and lead scalars if they are initially nonuniform. This can potentially lead to unphysical values of derived microphysical quantities from advection. To investigate this aspect, a simple extension of SFVT is tested to determine the practical
importance of ensuring monotonicity of the ratios of advected scalars for WRF-PD. While more sophisticated methods have been developed to preserve or nearly preserve monotonicity of scalar ratios (e.g., Schar and Smolarkiewicz 1996; Pfahl et al. 2012), a very simple approach was used here to estimate the effects of preserving monotonicity. In this approach, an operator is applied that adjusts the secondary scalar to ensure that the minimum and maximum values of the ratios of the secondary and lead scalar after advection cannot exceed values prior to advection within the grid stencil region of the lead scalar flux calculation. This operator is applied to the secondary scalars $Q_{rim}$, $B_{rim}$, $N_r$, and $N_r$. For simplicity it does not globally conserve secondary scalars if adjustments are applied, that is, it does not guarantee that the domain-total quantity of the secondary scalar is preserved during advection. However, while adjustments are not infrequent (typically 4%–8% of grid points with nonzero $Q_i$ at each time step), they are generally very small and, hence, the impact on global conservation is negligible. For example, the magnitude of the sum of $Q_{rim}$ adjustment at each grid point in a time step divided by the sum of $Q_{rim}$ at each grid point is generally $10^{-3}$–$10^{-6}$. Overall there is only a limited impact of applying this operator (e.g., cf. the green and red solid lines in Fig. 9), and, hence, preserving monotonicity of derived quantities does not appear to be particularly important here.

While solutions using SFVT are quite similar to those using the traditional approach, the computational cost is notably reduced (Table 8). Cost is calculated by taking an average of the total wall clock time per time step of three runs for each test (there is limited variability between individual runs for a given test). Runs were performed on the National Center for Atmospheric Research Yellowstone supercomputer on a single processor, using the default Linux x86_64 Intel ifort compiler. WRF-PD-5-SFVT and WENO-5-SFVT give a reduction of mean wall clock run time per time step by 10.7%–11.7% relative to WRF-PD and WENO, respectively. The reduction in cost is somewhat greater using WRF-PD-3-SFVT and WENO-3-SFVT, with a reduction of 11.3%–14.7% relative to WRF-PD and WENO. Thus, there is small increase in efficiency using $W^{(Third)}_{i+1/2}$, but the cost difference between $W^{(Third)}_{i+1/2}$ and $W^{(Fifth)}_{i+1/2}$ is small. Applying the monotonic adjustment operator with SFVT results in only a small increase in cost relative to WRF-PD-5-SFVT.

The primary reason for the cost reduction using SFVT is because calculating the linear $W$, analogous to the linear flux calculations in S08, is highly efficient, while the nonlinear PD limiter and flux calculations in WENO are expensive. Thus, WRF-S08 has a mean total wall clock time per time step smaller by 15.4% compared to WRF-PD, although as explained above it gives considerably different solutions. SFVT is somewhat less
efficient than WRF-S08 because several lead scalars are advected using the unmodified WRF-PD or WENO schemes (potential temperature, water vapor mixing ratio, $Q_c$, $Q_r$, and $Q_i$), but the approach reduces cost in run time by about 3% per secondary scalar and, hence, becomes more efficient as the ratio of the number of secondary to lead scalars increases.

Also shown in Table 8 is the computational cost of squall-line tests using either the one-moment WSM6 (Hong and Lim 2006) or the partially two-moment Thompson (Thompson et al. 2008) microphysics schemes in conjunction with WRF-PD advection using the traditional approach. The P3 microphysics using WRF-PD is somewhat faster than WSM6 or Thompson (~10% reduction in wall clock time). However, the cost differences are much greater using P3 with SFVT, with WSM6 and Thompson having a wall clock run time larger by 22%–24%, despite P3 having the same number of prognostic microphysics variables as Thompson and two more than WSM6. Some increase in efficiency is likely using Thompson with SFVT, although the cost reduction is expected to be modest because Thompson has five mass mixing ratio scalars, but only two number mixing ratio scalars, meaning there are only two secondary scalars versus five lead scalars. WSM6, or any one-moment scheme, cannot be used with SFVT since only mass mixing ratios are prognosed.

### 5. Summary and conclusions

In this paper, issues related to advection of coupled hydrometeor quantities in microphysics schemes are discussed. For physically consistent solutions, intensive microphysical quantities derived from advected prognostic variables in multimoment/multivariable bulk schemes should be preserved when the advected variables are linearly related to one another within a finite hydrometeor region [linear cloud advection (LCA)]. While it is well known that linear or semilinear advection schemes preserve linear relationships between advected variables (e.g., Thuburn and McIntyre 1997; Lauritzen and Thuburn 2012), it is less clear if quantities derived from the advected variables are preserved for LCA when these quantities are complicated functions of the advected variables. To address this issue, general criteria were derived for functional forms of quantities that are preserved during LCA using linear or semilinear schemes. Intensive quantities with functional forms that do not meet these criteria are not preserved for LCA, and, hence, physically inconsistent. This is

---

1 We refer to the Thompson scheme as “partially” two moment because in the version tested, it prognoses mass and number mixing ratios for cloud ice and rain, and mass mixing ratios only for cloud water, graupel, and snow.
useful for understanding the behavior of derived quantities, which may otherwise not be obvious. As an example, it was shown that an initially uniform PSD shape parameter $\mu$ in the three-moment bulk microphysics scheme of Milbrandt and Yau (2005), despite its complicated functional dependence on the three prognostic advected moment variables, meets these criteria and, hence, was preserved for LCA. However, an initially uniform $\mu$ was not preserved if at least two of the advected variables were not linearly related, in contrast to LCA. These results were supported by numerical tests.

A new flux-based approach to advect coupled microphysical quantities was proposed, which we call scaled flux vector transport (SFVT). In this approach, similar to the vector transport (VT) method applied previously in aerosol models, a lead microphysical scalar (mass mixing ratio) is chosen for each hydrometeor category and advected using the host model’s unmodified advection scheme. Advection of secondary microphysical scalars

![Figure 12](image_url)

**FIG. 12.** As in Fig. 5, but for the WRF-S08 test.

<table>
<thead>
<tr>
<th>Name</th>
<th>Mean wall clock time (s)</th>
<th>Percent change using SFVT compared to the corresponding test using the traditional approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRF-S08</td>
<td>0.176</td>
<td>-15.4%(^a)</td>
</tr>
<tr>
<td>WRF-PD</td>
<td>0.208</td>
<td></td>
</tr>
<tr>
<td>WRF-PD-5-SFVT</td>
<td>0.186</td>
<td>-10.7%</td>
</tr>
<tr>
<td>WRF-PD-3-SFVT</td>
<td>0.185</td>
<td>-11.2%</td>
</tr>
<tr>
<td>WRF-PD-5M-SFVT</td>
<td>0.188</td>
<td>-9.7%</td>
</tr>
<tr>
<td>WENO</td>
<td>0.222</td>
<td></td>
</tr>
<tr>
<td>WENO-5-SFVT</td>
<td>0.196</td>
<td>-11.7%</td>
</tr>
<tr>
<td>WENO-3-SFVT</td>
<td>0.190</td>
<td>-14.4%</td>
</tr>
<tr>
<td>WSM6</td>
<td>0.228</td>
<td>+22.6%(^b)</td>
</tr>
<tr>
<td>THOMPSON</td>
<td>0.230</td>
<td>+23.7%(^b)</td>
</tr>
</tbody>
</table>

\(^a\) Calculated relative to WRF-PD.

\(^b\) Calculated relative to WRF-PD-5-SFVT.
related to the lead scalar is calculated by scaling the lead scalar fluxes using linear weighting functions that map onto the grid stencil of the lead scalar fluxes. Because the weighting function calculations are much simpler than the nonlinear advective flux calculations used in models that retain positive definiteness and/or monotonicity, the main advantage of SFVT is its computational efficiency. The gain in efficiency compared to the traditional approach depends on the cost of the nonlinear advection calculations, with reduced benefit compared to using inexpensive methods such as hole filling. The efficiency gain may also be reduced if communication across grid tiles is a significant factor in the overall cost; if the same order weighting functions as the nonlinear flux calculations are used, then communication is essentially unchanged. SFVT is also simple to apply with any flux-based advection scheme, unlike VT, which requires reverse engineering of advection scheme code and can be difficult to implement. Idealized 1D tests showed that accuracy using SFVT is similar to the traditional approach of separately advecting each scalar, relative to benchmark analytic solutions. By design it also preserves linear relationships between advected scalars. SFVT does not preserve monotonicity of secondary scalars and, hence, does not conserve these scalars if adjustments are made to keep intensive quantities derived from them within certain ranges. However, tests using WRF showed that preserving monotonicity in this instance is unimportant.

SFVT was tested using the P3 microphysics scheme in WRF for an idealized two-dimensional squall-line case. Tests using two different advection schemes were described: the standard WRF scheme with a positive-definite flux limiter (WRF-PD), and fifth-order WENO. Lead scalars were chosen as the mass mixing ratio in each hydrometeor category of P3 (cloud water, rain, ice), and the secondary scalars were the rain number, ice number, rime mass, and rime volume mixing ratios. SFVT with either advection scheme provided solutions very similar to those using the traditional approach of independently advecting each scalar, relative to benchmark analytic solutions. By design it also preserves linear relationships between advected scalars. SFVT does not preserve monotonicity of secondary scalars and, hence, does not conserve these scalars if adjustments are made to keep intensive quantities derived from them within certain ranges. However, tests using WRF showed that preserving monotonicity in this instance is unimportant.

SFVT was tested using the P3 microphysics scheme in WRF for an idealized two-dimensional squall-line case. Tests using two different advection schemes were described: the standard WRF scheme with a positive-definite flux limiter (WRF-PD), and fifth-order WENO. Lead scalars were chosen as the mass mixing ratio in each hydrometeor category of P3 (cloud water, rain, ice), and the secondary scalars were the rain number, ice number, rime mass, and rime volume mixing ratios. SFVT with either advection scheme provided solutions very similar to those using the traditional approach of independently advecting each scalar, relative to benchmark analytic solutions. By design it also preserves linear relationships between advected scalars. SFVT does not preserve monotonicity of secondary scalars and, hence, does not conserve these scalars if adjustments are made to keep intensive quantities derived from them within certain ranges. However, tests using WRF showed that preserving monotonicity in this instance is unimportant.

Acknowledgments. This work was partially supported by U.S. DOE ASR DE-SC0008648, NASA NNX14AO85G, and the NSF Science and Technology Center for Multiscale Modeling of Atmospheric Processes (CMMAP), managed by Colorado State University under Cooperative Agreement ATM-0425247. J. Harrington was supported by the National Science Foundation under Grant AGS-1433201. A. Jensen and J. Harrington were supported by U.S. DOE ASR DE-SC0012827. The authors thank Dr. Peter Lauritzen for comments on an earlier version of the manuscript. Computing support was provided by NCAR’s Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

REFERENCES


