New Insights into Large Eddy Simulation

J. P. Boris, F. F. Grinstein, E. S. Oran, and R. L. Kolbe

Laboratory for Computational Physics and Fluid Dynamics

April 30, 1992
Fluid dynamic turbulence is one of the most challenging computational physics problems because of the extremely wide range of time and space scales involved, the strong nonlinearity of the government equations, and the many practical and important applications. While most linear fluid instabilities are well understood, the nonlinear interactions among them makes even the relatively simple limit of homogeneous isotropic turbulence difficult to treat physically, mathematically, and computationally. Turbulence is modeled computationally by a two-stage bootstrap process. The first stage, Direct Numerical Simulations, attempts to resolve the relevant physical time and space scales but its application is limited to diffusive flows with a relatively small Reynolds number ($R_e$). Using Direct Numerical Simulation to provide a database, in turn, allows calibration of phenomenological turbulence models for engineering applications. Large Eddy Simulation incorporates a form of turbulence modelling applicable when the large-scale flows of interest are intrinsically time dependent, thus throwing common statistical models into question. A promising approach to Large Eddy Simulation involves the use of high-resolution monotone computational fluid dynamics algorithms such as Flux-Corrected Transport or the Piecewise Parabolic Method which have intrinsic subgrid turbulence models coupled naturally to the resolves scales in the computed flow. The physical considerations underlying and evidence supporting this Monotone Integrated Large Eddy Simulation approach are discussed.
**CONTENTS**

1. INTRODUCTION .................................................................................................................. 1

2. BASIC CONCEPTS AND APPROACHES IN LARGE EDDY SIMULATION .................. 3

3. COMPUTATIONAL FLUID DYNAMICS REQUIREMENTS FOR DIRECT NUMERICAL SIMULATION AND LARGE EDDY SIMULATION .................................................. 10

4. EVIDENCE THAT MONOTONE ALGORITHMS HAVE BUILT-IN SUBGRID MODELS .......................................................... 17

5. CALIBRATING FLUX-CORRECTED TRANSPORT FOR LARGE EDDY SIMULATION ........................................... 25

6. A HYDRODYNAMIC ANALOGY FOR TURBULENCE MODELLING ............................. 33

7. SUMMARY AND DISCUSSION ................................................................................. 35

ACKNOWLEDGEMENTS ............................................................................................... 39

REFERENCES .............................................................................................................. 40
NEW INSIGHTS INTO LARGE EDDY SIMULATION

1. Introduction

This paper considers an approach to Large Eddy Simulation (LES) using built-in subgrid turbulence models which appear naturally from the monotone Computational Fluid Dynamics (CFD) algorithms used to simulate the resolved components of the flow. This approach differs somewhat from the conventional LES approaches reviewed, for example, by Reynolds [1], although much of the terminology and goals are the same: "to compute the three-dimensional time-dependent details of the largest scales of motion (those responsible for the primary transport) using a simple model for the smaller scales. LES is intended to be useful in the study of turbulence physics at high $R_e$, in the development of turbulence models, and for predicting flows of technical interest in demanding complex situations where simpler model approaches (e.g. Reynolds stress transport) are inadequate."

The differences between this Monotone Integrated Large Eddy Simulation (MILES) approach and conventional LES approaches are quite basic, however, and arise from how certain necessary tradeoffs are made and how best to optimize the overall performance of the CFD model. Evidence supporting this MILES viewpoint of LES subgrid modelling is based on the successful use of Flux-Corrected Transport (FCT) and other monotone algorithms for solving time-dependent CFD problems with steep gradients and turbulence and on certain common-sense considerations in making necessary numerical tradeoffs. It is only in the last several years, however, that a connection has been drawn between the grid-scale behavior of these algorithms and the need for and required properties of a subgrid-stress model to represent the unresolved "turbulent" scales.

Section 2 reviews the Basic Concepts and Approaches in Large-Eddy Simulation and includes a discussion of the desired properties of a good subgrid turbulence model. Section 3 presents a discussion of Computational Fluid Dynamics Requirements for Direct Numerical Simulation and Large Eddy Simulation, highlighting the close interaction between the grid-scale errors in the underlying CFD algorithm and the subgrid turbulence model. Section 4 is devoted to a general discussion of Evidence That Monotone Algorithms Have Built-In Subgrid Models. Some of the general aspects of the MILES viewpoint have been presented earlier [2–4]. Attention in this paper is centered around FCT algorithms because of our

experience using them at NRL, though the ideas and results should also apply to other monotone and effectively monotone CFD algorithms. Successes over the last two decades using FCT with no explicit turbulence model to simulate flows which are expected to be turbulent at small scales have been surprising. This body of direct and indirect evidence is also discussed in Section 4. In an effort to understand why these computer models are working as well as they are, we have studied what these models actually do to complex fluid flows which may be called turbulent and attempted to understand the intrinsically imprecise notions involved in LES.

Calibrating Flux-Corrected Transport for Large Eddy Simulation is considered in Section 5. Truly definitive numerical tests of LES are still difficult to define. Because of the limited range of resolved time and space scales available to direct simulation, the masking influence of physical viscosity is very strong in LES runs for which corresponding Navier-Stokes solutions are available. The dissipation provided by the viscosity reduces the cascade of small-scale structure to unresolved scales in the LES model being tested. Conversely, when a higher-resolution LES solution is used as the basis for determining convergence of another LES simulation, the short wavelengths in the inertial range are not strongly damped by the physics but there is always doubt as to whether the higher-resolution calculation is "correct." Further, analyzing theoretically exactly how given CFD algorithms will actually fit together with particular subgrid turbulence models is not practical and, as pointed out by Frisch [5], probably is not even possible.

Section 6 presents a qualitative way to understand these general aspects of LES through A Hydrodynamic Analogy for Turbulence Modelling. The cascade of energy from macroscopic scales through the inertial range of Kolmogorov to dissipation by viscosity is likened to water being poured into the center of a flat table and flowing smoothly off the edges. This analogy makes the interplay of the various scales in turbulence and cascade easier to interpret and leads naturally into the Summary and Discussion of Section 7.

The objective of this paper is to organize, quantify and at least partially substantiate the strong evidence suggesting that monotone convection algorithms, designed to satisfy the physical requirements of positivity and causality, have a minimal nonlinear LES filter and matching subgrid "turbulence" model already built in. The positivity and causality
properties of FCT and other monotone algorithms, properties not present in most commonly used convection algorithms, appear to ensure efficient transfer of the smallest grid-scale motions, generated by computationally resolved fluid dynamic mechanisms, smoothly off the resolved grid. This occurs with minimal contamination of the well-resolved scales. This conclusion, and the reasoning and computations which support it, are explained in the sections to follow.

2. Basic Concepts and Approaches in Large Eddy Simulation

Recent discussions of the usual approaches to LES can be found, for example, in articles by Reynolds [1], Hussaini and Speziale [6], Wyngard [7], Rogallo and Moin [8], and Speziale [9]. These authors reference earlier developments and concepts introduced in a number of papers including Smagorinsky [10], Lilly [11], Deardorff [12], Leonard [13], Bardina, Ferziger and Reynolds [14,15], and Biringen and Reynolds [16]. More recently, attention has turned to necessary extensions and generalizations such as compressible LES (e.g., papers by Speziale et al. [17] and Zang, Dahlburg and Dahlburg [18]), to considerations of models for stochastic backscatter from the unresolved scales into the resolved scales by Leith [19], to LES subgrid models which depend more correctly on local features of the flow (e.g., papers by Germano et al. [20] and Piomelli et al. [21]), and to the Monotone Integrated Large Eddy Simulation (MILES) models [2–4] which are the subject of this paper.

The usual approaches to LES methodology focus on the flow features which are large enough to be resolved by the CFD model. This is accomplished by selecting a filter function $F(x - x')$ which is convolved with a flow variable, $f(x, t)$, to define filtered or macroscopic variables:

$$\overline{f}(x, t) \equiv \int f(x', t) F(x - x') \, dx'.$$

By definition, therefore, these macroscopic, filtered variables have little or no short wavelength structure because it has been filtered out. The unknown short wavelength information, lost through the filtering, is called the residual or subgrid component $f'(x, t)$. If we knew $f'(x, t)$, the full, correct solution, $f(x, t)$, would be given by

$$f(x, t) \equiv \overline{f}(x, t) + f'(x, t).$$
Complete knowledge of quantities such as the mass density $\rho(x,t)$, the fluid velocity $u(x,t)$, and the pressure $P(x,t)$, resolved on all length scales down to the Kolmogorov scale, is beyond reasonable expectation except for flows with very low $R_e$. Indeed, we probably would be unable to deal with all the data computationally if it were available. LES, therefore, is founded on the reasonable expectation that macroscopic quantities of practical interest, such as turbulent mass, momentum, and energy transfer, drag, inter-species entrainment, depend primarily on the macroscopic variables $\overline{\rho}(x,t), \overline{u}(x,t), \overline{P}(x,t)$, which can be determined computationally with adequate accuracy. Any residual dependence of macroscopic or averaged quantities on the unresolved subgrid component of the fluid dynamic variables, it is argued, can be modelled by simple expressions involving only the computed macroscopic quantities.

The incompressible Navier-Stokes equations can be filtered using equation (1) in the same way as the individual fluid variables when the filter integral commutes with the partial derivatives. The results,

$$\frac{\partial \overline{u}_i}{\partial x_i} = 0 \quad \text{and}$$

$$\frac{\partial \overline{u}_i}{\partial t} + \frac{\partial \overline{u}_i \overline{u}_j}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \overline{P}}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + \nu \nabla^2 \overline{u}_i$$

are well known (see for example, [1, 6, 8, 20]). In equation (4) the subgrid stress tensor $\tau_{ij}$, containing the unknown information from the residual or subgrid fluid velocities, is defined as

$$\tau_{ij} \equiv \overline{u}_i \overline{u}_j - \overline{u}_i \overline{u}_j. \quad (5a)$$

Alternately, the subgrid stress tensor can be written explicitly in terms of the residual and resolved velocity components,

$$\tau_{ij} = \overline{u}_i \overline{u}_j + \overline{u}_i u_j' + \overline{u}_j u_i' + u_i' u_j' - \overline{u}_i \overline{u}_j. \quad (5b)$$

The conventional model of this subgrid stress is the Smagorinsky model,

$$\tau_{ij} = -2 \nu_S S_{ij}, \quad (6)$$
where
\[ \nu_e \equiv (K_s \Delta)^2 \left( 2 \, \overline{S}_{ij} \, \overline{S}_{ij} \right)^{\frac{1}{2}}. \] (7)

Here \( \Delta \) is taken as the cell size, \( K_s \) is the Smagorinsky constant, and \( \nu_e \) is the eddy viscosity coefficient. Juggling these constants near walls and for other special conditions to improve the performance of the overall LES model has become an art form. In eqs. (6) and (7), \( \overline{S}_{ij} \) is the strain-rate of the macroscopic flow,
\[ \overline{S}_{ij} \equiv \frac{1}{2} \left[ \frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right]. \] (8)

Because of the divergence term, \( \partial \tau_{ij}/\partial x_j \), in equation (4), the principal effect of modelling the subgrid stress in this way is to add diffusion to the macroscopic equations to represent the cumulative effects of the unresolved small scales.

Hussaini and Speziale [6], referring to this class of LES approach, noted that “there are some major difficulties with LES that need to be overcome before it can yield reliable and economically feasible predictions for the complex turbulent flows of scientific and engineering interest. These problems are as follows:

(i) the implementation of LES in spectral domain decomposition or high-order finite difference codes so that complex geometries can be treated;

(ii) the development of improved subgrid scale models for strongly inhomogeneous turbulent flows (e.g., flows with localized regions of relaminarization or large mean velocity gradients);

(iii) the development of reliable a priori tests for the screening of new subgrid scale models;

(vi) the problem of defiltering;

(v) the problem of modifying subgrid scale models to accommodate integrations to solid boundaries.”

The fact that the mathematical procedures leading to equations (3) and (4) are invalid when the filter operator does not commute with the time and space derivatives (as occurs with variable grids or when it is advantageous to use a space- or time-varying filter) is
usually overlooked. This may not be a serious drawback. Since the undetermined subgrid stress tensor must be modelled by a phenomenology in any case, the necessary terms to correct these mathematical consistency problems can also be lumped into the subgrid phenomenology. LES approaches using the Smagorinsky approximation, based in part on an assumed isotropy of the unresolved small scales, replace the divergence of the subgrid stress tensor with a conservative divergence of diffusive fluxes based on an eddy viscosity, whose strength depends on gradients of the macroscopic velocities with one or more adjustable coefficients. This eddy viscosity is the subgrid turbulence phenomenology which approximates the effects of unresolved small-scale motions on the resolved fluid dynamic variables.

In some cases, the filter function is a spatial average over a volume in the fluid corresponding to one or more cells of the computational domain (see, for example, [1,16,20]). In models based on spectral representations, this filtering is done in k-space and a cutoff is applied to short wavelengths directly in the Fourier spectrum; in some cases, a sharp cutoff is used and in others, a Gaussian is used to reduce short wavelength components smoothly [16]. A drawback to these various approaches is the relatively large amount of smoothing (filtering) needed to ensure that the short wavelength content remaining in the computed solutions does not contribute significantly to their error. This filtering, in some formulations, can have a significant dissipative effect even on scales that can be resolved well. The eddy viscosity also tends to increase the dissipation and can inhibit the linear growth of instabilities in laminar fluid systems whose gradients are improperly interpreted as sources of eddy viscosity. The Smagorinsky model can lead [21] “to the decay of the perturbations even in the instances in which the flow should have been unstable.”

As can be seen for the example of the incompressible Navier-Stokes equations above, the additional terms arise from applying the filter to products of the unfiltered variables, i.e., from the nonlinear terms. In compressible LES, there are of course, many more nonlinear terms to be considered. In addition to determining how or whether to model all these nonlinearities, the user of LES also must prescribe how to convert the answers computed using smoothed (filtered) variables back to the original unsmoothed (defiltered) representation [6].
The ideas underlying these LES approaches are all relatively natural and emphasize the pivotal role accuracy plays. The solutions of the filtered equations contain little or no short wavelength structure even when the high $R_e$ flow being simulated does. Therefore it is reasonable to assume that the filtered solutions can be determined accurately using a discrete computational representation (either a grid or a finite spectral representation), whereas the correct solution to the unfiltered equations cannot be found so easily numerically. The tradeoff requirement to model the filtered products of the unknown (unfiltered) solutions, the so-called subgrid turbulence model, is considered the lesser of the two evils. By filtering the variables and the equations, accurate solution of the resulting system can be expected. The control of numerical error is still very important in standard LES models, however. Although the choice of filter can be used to reduce the short-wavelength content of the numerical solutions, the short wavelengths generally cannot be completely removed. Further, the price paid to ensure that the inaccurate short wavelengths can be controlled is the modification of the longer-wavelength structures which could otherwise be computed more accurately.

Several assertions about modelling small scales in fluid dynamics seem obvious but perhaps should be stated. If the fluid dynamic interactions at any particular scale are not accurately resolved, no numerical model can give more than an approximation to the true behavior of the flow at that scale. Further, no computational model is perfect. Therefore, any CFD simulation model will contain imperfections arising from the necessary tradeoffs that have to be made to construct it. These two assertions about the process of modelling the fluid dynamics are followed by two assertions about the fluid dynamics. The effect of each small unresolved fluid dynamic structure on the resolved macroscopic properties of the flow is small though the composite, integrated effects can be appreciable. In this regard, the largest unresolved scales are generally most important; progressively smaller unresolved scales contribute less and less to the large-scale behavior.

These assertions lead to the conclusion that a range of scales as wide as possible should be directly simulated; the subgrid models in LES should be restricted to the unresolved scales to the greatest extent possible with minimal effect on the resolved scales. On the one hand, accurately calculating as much as possible has to be a good idea. On the other
hand, the second two assertions suggest that the fluid does not strongly conspire against
the necessary LES partitioning into resolved and unresolved subgrid fields. Indeed, it has
been pointed out, as an example of a fundamental fluid dynamic result obtained from DNS,
that nearby scales seem to interact most strongly [1]. The main effect of the large scales
on widely separated small scales seems to be vortex stretching arising in the mean strain
fields.

An ideal subgrid model should have the following properties:

P1. It should apply without restriction to the fluid dynamic model being solved macro-
scopically, e.g., it should handle compressibility and high Mach number flow, multispecies
effects, etc., as appropriate to the problem at hand.

P2. It should satisfy the global conservation laws of the system as integrated over the
resolved and unresolved scales.

P3. It should minimize the contamination of macroscopic scales by the inaccurately re-
solved flow structures on the grid scale and by the numerical filtering. This allows the
resolvable linear and nonlinear processes which physically drive the subgrid dynamics to
be calculated as accurately as possible.

P4. It should accomplish the physical mixing and averaging expected of the complex but
unresolved flows on the correct macroscopic space and timescales.

P5. It should smoothly connect to the resolved macroscale solutions at each point in space,
even for variable grid size. The effects of all scale lengths, whether modeled or resolved,
should be included exactly once.

In addition, several conditions have to be satisfied in the high $R_e$ fluid dynamic system
being modeled and the set of equations being used to ensure that the LES–subgrid model
approach makes sense. These conditions are based in part on the self evident assertions
made above and in part on distinctions between LES and Very Large Eddy Simulation
(VLES) as introduced in [1]. These conditions are:

1. The problem being solved is such that the macroscopic LES model can resolve the
dynamics of the energy containing, turbulence-driving scales,

2. The macroscopic convection velocities are sufficiently larger than the unresolved turbulence velocities that small-scale turbulent motion of material, mass, momentum, and energy accounts for a small portion of the global transport in the problem, and

3. Unresolved "turbulent" diffusion dominates molecular transport or else the molecular effects are explicitly included in the LES model equations.

Without these three conditions being satisfied, the expectation of any subgrid turbulence model working is small. Fortunately, any system being tackled by LES satisfies these conditions essentially by definition. Conditions 1 and 2 above guarantee that the resolved component of the fluid dynamics which is treated accurately contains most of the information of interest. Were this not the case, most of the solution would depend on the subgrid model, casting the predictive capability into the realm of phenomenology. Condition 3 says that any transport phenomenon which is not resolved convection, and that is at least as important as the unresolved convection, is also included in the macroscopic model.

Dynamic subgrid models of turbulence are being developed and tested as improvements to this general approach. Germano et al. [20] recently proposed a subgrid stress model attempting to overcome the deficiencies "by locally calculating the eddy viscosity coefficient to reflect closely the state of the flow. This is done by sampling the smallest resolved scales and using this information to model the subgrid scales." This sampling is accomplished by using a second filter called the "test" filter which is broader than the underlying LES "grid" filter. The difference between the two subgrid scale stress tensors calculated using these filters is called the "resolved turbulent stress" by these authors who then use the Smagorinsky model to relate the just resolved large scale viscosity derivatives to the eddy viscosity.

This approach has several advantages. In laminar flow or at solid boundaries, the difference between the two filtered stress approximations should vanish. Thus additional wall damping functions or other phenomenologies are claimed to be unnecessary. Further, since the average dissipation of the model can be of either sign, this "dynamic subgrid-scale eddy viscosity" model apparently does not rule out deterministic backscatter. The
remaining disadvantages appear to be related to the differencing procedure used to identify the small scale (but resolved) turbulent stresses. Because a difference is taken which must subsequently be divided out, the expression for the local time-dependent "dynamic" Smagorinsky coefficient can blow up. The spatial average taken over planes in the flow to remove this singularity also reduces the desired locality of the model.

Another assumption in references [20] and [21] is the use of the Smagorinsky model as the functional expression of the estimated turbulence stress closure at both filter scales. This has been a good starting point, allowing the cancellations which are claimed to give the model its advantageous properties near walls. The adjustable parameter in the model is the ratio of the two filter scales, assuming that the smaller filter is the grid scale. Tradeoffs are involved here. Further, the backscatter allowed in the model is only deterministic as no stochastic component is postulated. Therefore the important role of small-scale structures in the flow triggering large scale instability is inhibited [19,21]. We will return to this issue of backscatter in the next section where we discuss the extent to which monotone, flux-limiting convection algorithms such as FCT contain an adequate built-in filter and subgrid model.

3. Computational Fluid Dynamics Requirements for Direct Numerical Simulation and Large Eddy Simulation

Though accuracy is extremely important, successful large simulations must usually trade some accuracy for increased efficiency, flexibility, and generality. For example, DNS problems have generally been treated by spectral algorithms because of their high accuracy in well-resolved wavelength regimes. Rai and Moin [22], however, recently used finite differences effectively in DNS of the Navier-Stokes equations because of their relative efficiency. Therefore it should not be surprising that other CFD algorithms should be explored for LES as well.

The CFD requirements for DNS and LES are, in fact, different. In DNS the smallest resolved scales are continuously being smoothed and dissipated by viscosity. The relative motions at these scales are quite slow so the amplitudes of the highest harmonics of the
corresponding field variables are small. Local numerical errors in the short wavelengths can have little effect. Since spectral methods excel at intermediate and long wavelength where physical viscosity gives relatively little smoothing, they generally have been a good match for DNS problems. In LES, the Reynolds number of the flow one wishes to treat is so large that viscosity is not effective in removing steep gradients on the smallest resolved scales. The spectral energy content of motions and gradients on these scales is thus correspondingly larger in LES problems than in DNS problems, in complete accord with the relatively slow decrease of energy content with wavenumber in the inertial subrange of a Kolmogorov spectrum [eg., 23–25]. It has been known for some time (Leonard [13]) that “Modifications of the Navier-Stokes equations must be introduced to simulate properly the energy cascade. Considerable “damming up” of the turbulence energy in the large scales would occur, for example, if the unmodified equations were used with an energy-conserving finite-difference scheme on the advective term.”

The view of LES expressed here differs from that in [1] since we believe it is not practical “to separate the formulation of the LES problem from the numerical method used for its solution.” Such a separation is attractive as it enables numerical analysis of the resulting methods to parallel Reynolds stress analysis. Unfortunately the dividends from this analysis are unsatisfying and incomplete because closure problems remain. Furthermore, defiltering the resolved-field solutions to obtain information about the physically meaningful fields is a nuisance and the ad hoc subgrid stress models require empirical calibration by experiments and simulations. Unless LES methodology with strong filtering is used, the “subgrid fields” have to be matched to the “resolved fields” at the smallest resolved scales - just where the distinctions between various methods and algorithms are greatest and numerical errors are largest. Since this matching should be done with some representation of the fluid dynamics at all scales included once but not twice, the short-wavelength errors of the CFD algorithm should not be ignored in developing a subgrid turbulence model designed to incorporate physics occurring at the grid scale and below into the flow equations governing the large-scales.

By filtering the mathematical model in the usual way to obtain LES equations at the resolved scales, sufficient smoothing is added so that the otherwise underresolved Navier-
Stokes equations will be well behaved at the grid scale – even for conventional algorithms not designed to control gradients at this scale. The price is a rather substantial influence of the filtering at larger scales where most algorithms would be accurate, even on the unfiltered equations. Conventional “subgrid” models are nominally formulated independently of the CFD algorithm being used but need to take the effects of the LES filtering on the well-resolved scales specifically into account, effectively extending the phenomenological modeling far into the longer scale lengths where it should not be needed.

Reynolds [1] also discussed the notion of LES performed with no subgrid models, but did not address the interpretation that these LES systems may already have a minimal built-in subgrid model, particularly if they are effectively monotone. His discussion of this notion was addressed primarily to the high-$Re$ simulations being reported in reference [26] in which the resolution in free space was inadequate to resolve the vortices being shed from highly-resolved boundary layers. The third-order upwind algorithm used by Kuwahara and co-workers [26-28] is not monotone in the strict sense but has a fourth order dissipation which apparently is strong enough to channel the grid scale fluctuations smoothly into the unresolved inertial range as required of a functioning LES subgrid model without using an explicit filter. In MILES models, as in all others LES models, residual $Re$-dependent effects of course cannot be simulated without viscosity appearing explicitly or through a phenomenology. Further, boundary layer phenomenologies are still needed when wall regions are underresolved. Reynolds observes that “LES is very resilient to the residual turbulence model.” Carrying this further, one can expect that a factor of two increase in the spatial resolution of adequately resolved LES and MILES models will bring more improvement in the fidelity of the well-resolved scales than an arbitrarily complicated subgrid model.

The FCT models used in most of the studies reported in Section 4 and for the simulations presented in Section 5 solve the continuity equations for mass, momentum, energy and any chemical species written in conservation form (see, for example, [29-32]):

\[ \frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0, \]  
\[ \frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{uu}) = - \nabla P, \]
\[
\frac{\partial E}{\partial t} + \nabla \cdot E \mathbf{u} = - \nabla \cdot P \mathbf{u}, \quad \text{and}
\]
\[
\frac{\partial n_i}{\partial t} + \nabla \cdot n_i \mathbf{u} = \left[ \frac{\partial n_i}{\partial t} \right]_{\text{chemistry}} \quad \text{for } i = 1, 2, ..., N_{\text{species}}.
\]

In the case of reactive shear flows [29,33,34], other physical processes such as diffusive transport and chemistry are coupled to convective transport using timestep splitting [e.g., 35]. The equations for convective transport were solved in most cases using a one-dimensional fourth-order phase-accurate FCT algorithm [36], directional timestep-splitting techniques on structured grids, and appropriate inflow and outflow boundary conditions [31,37]. Other more recent simulations [38-40], used multidimensional FCT algorithms, typically with fourth-order accuracy both in phase and amplitude.

In FCT the fluxes of mass, momentum, and energy between each resolved cell and its neighbors is calculated using a high order algorithm. These fluxes, to be completely consistent with the discrete finite-volume representation, are averages over the appropriate cell interface areas for a time interval corresponding to the timestep. Nevertheless, these fluxes contain some short wavelength information which cannot be properly resolved by the grid. Even with the high order or "infinite order" fluxes defined by spectral methods, the finite resolution of the discrete representation has an associated Gibbs phenomenon which accounts directly for nonphysical fluctuations of the order of 15% in numerically convected fluid dynamic quantities. These Gibbs fluctuations are absent only when the fluid profiles being convected are sufficiently smooth, i.e. have been filtered adequately.

If the intercell fluxes were known exactly, the cell values computed from summing these fluxes over all faces of the cell and updating the cell values accordingly would be exact. The flux-correction procedure uses what information is available about the errors in phase, amplitude and resolution of the composite numerical solution to limit the form these errors can take in the resolved-scale solution. This nonlinear correction flux appears as a reduction of a numerically calculated flux to a level determined to be consistent with the details of the fluid profiles and the grid resolution. This correction usually takes the form of an intermittent and local diffusion but backscatter is allowed and sometimes necessary to preserve the monotonicity/positivity properties of the real convected quantities [41]. In the momentum equations this is quite similar to the explicit addition of an eddy viscosity keyed
directly into the local instantaneous resolution and accuracy limitations of the underlying convection algorithm. This effective diffusivity can backscatter because it does not always appear as a positive diffusion. Further, it automatically counters the nonlinear effects causing the resolution problems in the first place.

For example, compression, shear, and vortex stretching all can generate unresolved short wavelength structure in the flow where it did not exist before. These are all velocity gradient effects which FCT detects as overlarge convective fluxes at some cell interfaces. These fluxes are limited as needed, effectively adding local or intermittent dissipation. This limiting or "correction" procedure would have no effect at all if the driving velocity gradients were not present. Thus the velocity gradients of the resolved-scales, with emphasis on the just barely resolved scales, lead directly to the nonlinear numerical filtering, as in conventional LES based on Smagorinsky-like models. In FCT models, however, velocity gradients in laminar, well-resolved regions do not lead to eddy transport so linear instabilities, e.g., [42-44], at moderate wavelengths are not hampered.

The residual numerical dissipation of the FCT algorithm in unsteady fluid simulations has been the subject of detailed investigations [41,45]. In the case of the simulation of free mixing layers, a small (second-order) numerical diffusion left in the model was investigated in the low-Mach-number regime [45]. Measurement of a global residual numerical diffusion was performed on uniform grids by comparing the laminar spread of the simulated mixing layer with that predicted by incompressible boundary layer theory. It was shown that the small residual numerical viscosity of the FCT algorithm, if left in the model, can emulate physical viscosity for laminar shear flows at moderately high $Re$. Based on these studies, modified FCT algorithms are being tested which push this residual diffusion to even higher order [45]. The global numerical diffusion was found to be essentially insensitive to changes in free-stream velocity ratio, and could be reduced rapidly in a predictable way by refining the grid spacing.

Since monotone convection algorithms were designed to limit errors in the shortest resolved scales in a physically meaningful way where sensible connection to a subgrid model is also required, they seem a better choice than linear convection algorithms for use in LES. Numerical experience at NRL and elsewhere (Section 4, below) suggests that the nonlinear
filter built into monotone CFD algorithms really serves the same purposes as a subgrid stress model. These MILES algorithms are derived from the physical laws of causality and positivity which also underpin convection in turbulent flows. They do minimal damage to the longer wavelengths while incorporating most of the local and global effects of the unresolved turbulence expected of LES subgrid models.

The tradeoff for satisfying positivity and causality and for the enhanced accuracy of monotone methods at short wavelengths is somewhat larger errors at long wavelengths than found in spectral methods. Since these long wavelength errors are very small in any case, the comparative advantage shifts to the monotone methods when accurate treatment of the smallest resolved scales is of paramount importance. These MILES algorithms work on and transform the fluctuations in the fluid field variables that cascade to short wavelengths due to resolved field nonlinear effects and instabilities. These unresolvable fluctuations are converted to the correct macroscopic variables locally but the timescale for this to occur is controlled by the resolved flow and not by microscopic physics. For example, viscous dissipation of the unresolved scales appears as heat since total energy is conserved and grid-scale kinetic energy is dissipated to maintain monotonicity. Diffusion of the eddy transport type is automatically left in the flow as required, but the fluctuating driving effects of random-phase, unresolved eddies on the large scales is missing unless specifically included as a subgrid phenomenology. This deficiency, however, is common in conventional subgrid models.

Figure 1 illustrates how a macroscopic quantity like entrainment, as defined in Section 5 below, is expected to behave as a function of computational resolution for a conventional LES model and for a MILES model in a system with turbulence. With linearly filtered LES algorithms and an explicit eddy viscosity, the correct entrainment for a turbulent high-$Re$ flow appears to be approached from above as shown by the upper curve. The computed solutions at finite resolution must be defiltered to correct some macroscopic quantities like the entrainment or the turbulent kinetic energy to their infinite numerical resolution values and defiltering is generally unstable.

With MILES algorithms, the effective filtering is nonlinear and thus the nonphysical diffusion does not extend significantly to long wavelengths. The curve labelled "mono-
tone algorithm" in the figure and marked with boxes to show changes by factors of two in resolution corresponds to a macroscopic quantity, here the volume of mixed and jet fluid, measured at a given time in a series of different-resolution simulations. This curve illustrates the minimum that is postulated to be found [3-4] at intermediate resolution. The minimum (extremum) is expected to occur when short wavelengths, which provide some entrainment, cannot be resolved but when the residual numerical diffusion present is smaller than the eddy diffusivity of the turbulent flow. This tradeoff is illustrated schematically by the shaded region in the lower right shown increasing transport (mixing) due to resolved eddies as the resolution is increased and the curve in the lower left showing increasing "transport" from the flux limiter in monotone algorithms as resolution is decreased. Beyond this minimum, increasing resolution actually increases the entrainment because removing the remaining grid-scale numerical filtering has less effect than adding the corresponding eddy diffusivity from the unresolved scales.

Monotone algorithms are generally devised using minimum dissipation to maintain monotonicity. If less diffusion is required to get the correct answer, the only solution is to increase the spatial resolution. Using a CFD algorithm with less dissipation is either unstable or leads, through blocking or damming up the cascading short wavelength structure, to nonphysical solutions. FCT algorithms have been analyzed theoretically even though they are intrinsically nonlinear, and have been shown [46] to converge with increasing resolution to the correct solution of the underlying continuity equation being solved. This means that increasing resolution, even without any added subgrid transport model, will lead to a converged solution to the target high-$Re$ problem once the residual, resolution-based numerical dissipation has become smaller than the eddy diffusivity from all unresolved scales. This occurs once the grid spacing $\delta x$ is finer (smaller) than a critical value as shown schematically in Figure 1 at $log_2(R_{jet}/\delta x) = 3.0$. 

16
4. Evidence That Monotone Algorithms Have Built-In Subgrid Models

Our experience base for this paper is derived from use of the monotone FCT algorithms described above (e.g., [29,36,40,47-50,57]). Below, the acronym FCT will occasionally be used when it should be understood that the comments apply equally to a number of other monotone methods. Indeed, many comparable monotone methods now exist; see, Van Leer [51], Woodward and Colella [52,53], and the extensive references in [29] for examples. Three-dimensional LES using these methods has been performed for a number of problems with a focus on unsteady and highly transient systems.

The experience solving compressible flow problems, summarized below for FCT models, and related computational studies by others, e.g., Kuwahara and co-workers [26-28] and Woodward and co-workers [54-56], indicates that monotone CFD methods can actually be viewed as LES models with an intrinsic subgrid algorithm. This built in subgrid algorithm arises naturally from the nonlinear monotonicity-preserving ("flux-correction" or "flux-limiting") feature of these methods as described above in Section 3. FCT techniques were actually developed to treat unresolvable short wavelengths arising in nonlinear convection with no distinction between compressional, rotational, and potential aspects of the flow. Thus, it should come as no surprise that FCT is a good LES algorithm for turbulent flows although the historical use of FCT and other monotone algorithms has been primarily to simulate compressive (shock) phenomena.

In the next few paragraphs the use of FCT for time-dependent simulations of shear flows, jet flows, compressible turbulence, and Rayleigh-Taylor mixing will be reviewed. Originally these models were applied to problems with strong shocks and blast waves [52,53,57-60], detailed acoustic-vortex interaction studies [61,62], reactive shocks and detonation cell structure [60,63], and to other chemically reactive flow problems (see, for example, [29,64,65] and the references therein). Recently our FCT applications have migrated into the compressible shear flow and turbulence arenas [31,61,66-69] so that detailed comparisons in the usual DNS and LES contexts will begin to be available. The generally good agreement with experiments, other simulations, and known analytic solutions, where available, lends credence to the notion that FCT models are LES models without addition of an external subgrid turbulence model.

Extensive numerical simulations of subsonic, spatially evolving two- and three-dimen-
sional shear flows using FCT models have been performed by Grinstein and co-workers [31–34,37,38,71–77]. This work principally examines the evolution of large-scale coherent structures in the transitional regime within a few tens of diameters of the nozzle or splitter plate. The high-$Re$ experiments on these flows are turbulent throughout much of this regime however but the spatially-evolving simulations have been successfully compared with them despite the inability to resolve the small-scale structure and with no explicit subgrid scale model of its effects. The comparisons include both, instantaneous and time-averaged results. Close agreement with experimental observations was even found in two-dimensional simulations, including the asymmetric entrainment [71] and spreading rates [32,72] in a mixing layer, the distribution of quasi-stable vortex-pairing locations in self-excited circular jets [73]. Close agreement with high-$Re$ experiments was also obtained for comparisons of base-pressures and vortex-shedding frequencies in bluff-body near-wakes [74].

New fluid-mechanical information was obtained on global instabilities due to upstream feedback in free mixing layers [75,76], on the vortex-ring dynamics in circular jets leading to momentum-flux increases and negative turbulence-production [32], on mechanisms for passive pressure-drag control in the plane wake [74], and on the effects of chemical exothermicity on the shear-flow development [33,34]. The generally accurate prediction of the high-$Re$ flows being modelled would not have been possible if the basic FCT models – without any added turbulence model – did not have an effective subgrid model built in.

Moreover, very good agreement was also found when conducting (as possible) the more difficult comparisons based on three-dimensional laboratory and simulation data. The computations were shown to be capable of simulating the basic three-dimensional coherent-structure dynamics in transitional shear flows, including that of interacting spanwise rollers and horseshoe vortices in plane mixing layers [72], vortex reconnection leading to the formation of vortex loops in plane wakes [77], and axis switching and vortex-ring pairing in square jets [38]. While these simulations again focussed on large-scale dynamics, the real systems being modelled also have significant small-scale structure which the FCT models could not resolve. More recently [78], the effects of small-scale dynamics in the transitional shear flows was also investigated in addition to that of the large scales. The upstream turbulence present in laboratory plane-wakes was simulated by breaking the two-dimensional coherence of the inflow, perturbing the inflowing free-stream velocities
with suitable superpositions of sinusoidal modes with random phases and amplitudes. The validity of the approach was established by comparing the incoherent quantities resulting from ensemble averaging (eduction) of the three-dimensional coherent structures with those educed in laboratory experiments [79].

These models were also used in an extensive series of computations aimed at understanding and quantifying the generation of turbulence and the turbulent mixing of shock- and beam-heated channels in air such as arise in the propagation of lightning, lasers, and charged-particle beams. Experiments, e.g., [80,81] and references therein, showed cooling of the heated channels which could not be explained by molecular thermal conduction or the onset of convection. The Reynolds number of the resulting turbulence appeared to be very high. The theory developed to explain this was based on turbulent vorticity generation due to shock-density gradient interactions caused by asymmetries in the channel heating. This compressible theory was illustrated by FCT simulations which were subsequently calibrated using this theory and experiments [32,33,82,83]. The simulations were carried out for a long time to evaluate the turbulent growth and cooling of the channels in a number of circumstances including configurations where the asymmetries are intrinsically three-dimensional. The effective turbulent mixing diffusivities of the simulations agreed well with experiment, providing ample evidence that the underlying FCT models could deal with eddy thermal conductivity as well as eddy viscosity.

The generation of complicated mixing flows by strong Rayleigh-Taylor instabilities resulting from laser acceleration of thin foils is yet another class of compressible fluid dynamics problem considered extensively and successfully by FCT algorithms. Simulations of linear Rayleigh-Taylor growth with ablation [84,87,89], nonlinear mode saturation and foil disruption [85,86,88], and comparison with laboratory experiments [86–88,90] have all been performed to determine stability and symmetry requirements for laser-driven Inertial Confinement Fusion. These simulations include nonlinear thermal electron conduction and, in some cases, radiation transport to drive the instabilities in high-$R_e$ fluids where short scales are certainly excited by laser asymmetries and target imperfections. This is a physically more complex situation than ideal gas dynamics or hydrodynamics but the underlying fluid dynamics occurs with large perturbations and $R_e$ is usually in the range $10^4$ to $10^6$ where there is every expectation that a full spectrum of compressible turbulence will be excited.
These FCT simulations, used to support and guide the NRL experimental program for a number of years, have been compared in some detail with theory and experiment for high $Re$ flows with generally excellent agreement. The comparison of linear and non-linear growth behavior with theory and with experiments has been carried out at least as extensively for this type of flow as for the shear and jet flows reported above. Most of these computations were conducted in two dimensions but some calculations have also been done in three dimensions [89]. During the late stages of some of these experiments the flows are certainly turbulent on the small scales. If the underlying FCT algorithms were not accounting for the unresolved small scale motions in the fluid, at least approximately correctly, the quality of these comparisons could not have failed to show some major inconsistencies.

Other monotone algorithms in addition to FCT have been applied to high-$Re$ flows which physically are turbulent. The predominant use of these methods, however, has been for dynamic problems with shocks which either do not last long enough to teach us much about the intrinsic subgrid behavior or which are run on such inhomogeneous problems that the focus has been on the large-scale behavior. An exception is the work of Woodward and co-workers [54–56], who have looked at homogeneous high Mach-number turbulence with the Piecewise Parabolic Method (PPM) in much the same way that low Mach-number isotropic, homogeneous turbulence has been studied. Developed by Woodward and Colella [52,53], PPM is completely monotone, has been studied carefully, and has been optimized for a number of parallel and vector processing computers.

Using PPM one finds convergence of Euler (MILES) computations of increasing resolution to the solution obtained by very high resolution Navier-Stokes computations of the identical physical problem. The Kolomogorov $k^{-5/3}$ spectrum, really expected only as a transient because the flow is decaying rather than being driven at long wavelength, is seen as an envelope to the series of MILES spectra obtained at increasing resolution. This behavior has been seen for several two-dimensional configurations and the same behavior is also seen in three dimensions [56]. The converging spectra in this work, like the converging measures of entrainment soon to be discussed in Section 5, is rather direct confirmation of the existence and essential correctness of the effective subgrid model provided by the nonlinear flux-limiter in monotone algorithms.

Another group of researchers led by Kuwahara, e.g. [26–28], has had extensive experi-
ence in very low Mach number, turbulence-related simulations using a third-order upwind method which seems to be nearly monotone in that it has a fourth-order dissipation term which apparently obviates the need for additional linear damping (Section 3). This algorithm has been used without an explicit LES filter for extensive vortex shedding and vortex separation computations in both two and three dimensions for problems where the fluid has dynamic structure at scales that cannot be resolved even by the very fine grids employed. Small-scale vortices, generated in a well-resolved boundary layer, are convected into regions of the grid where they cannot be well-resolved. The model automatically filters these unresolvable fluctuations, apparently without damaging effects on the solution or on large-scale measures taken from the solution. Reynolds numbers as high as $10^6$ have been simulation this way, allowing the drag crisis on a circular cylinder to be demonstrated explicitly without an boundary phenomenology of any sort.

The observed ability of the FCT-based models to simulate the transitional shear-flow dynamics and post transition turbulent transport strongly supports the idea that the effective numerical dissipation due to the nonlinear FCT high-frequency filtering – combined with the conservative, causal and monotone properties of the algorithm, play the role of a a minimal subgrid model in these applications. Some of the monotone methods were specifically developed for shock problems, and, because they do such a good job of shock capturing, the bias still is toward high-speed applications. This usage pattern, coupled with the origin of many of the monotone methods naturally has led to the impression that all monotone methods are limited to high Mach number. This impression is incorrect for FCT and many of the other monotone algorithms.

When the physical viscosity is small, i.e. for underresolved Navier-Stokes flows, or alternately, when the computational cells are large, the large-scale features of solutions of the Navier-Stokes equations and the Euler equations are essentially identical using FCT. Both solutions show the effects of the flux-correction procedure as a residual, nonlinear filtering of short wavelengths. This filtering influences long wavelengths negligibly and yet is strong enough at short wavelength to prevent aliasing of high frequencies into the long wavelengths. This was shown for Berger’s equation [91] in detailed comparisons with a spectral model. It has subsequently been checked repeatedly for fluid dynamics through spatial convergence tests in every major configuration where FCT has been applied to jets, shear layers, and reacting flows, e.g. [32-34,61,76]. These tests have also been done in
three dimensions (see, e.g. Section 5 below) and generally show a converged long wavelength behavior when the system size is large enough to support at least a modest ratio between the energy containing long wavelengths and the eddies of a few cells wavelength which dissipate quickly. In the context of shear-flow simulations, the correct initial (linear) Kelvin-Helmholtz growth is ensured when the wavelength of the most amplified mode is resolved with 20 or more cells [44]. Consistent with this result it is found that coherent vortex structures more than 15-20 cells across change negligibly when the resolution is doubled or quadrupled to allow resolution of much smaller scales.

Any algorithm, including monotone algorithms, that uses knowledge of the grid relative to variations of the evolving solution, cannot be expected to be Galilean invariant. Adding a constant velocity to the flow everywhere moves real structures in the computed solution to different locations relative to the grid at the end of each timestep and they will be resolved differently if they have any grid-scale structure. Indeed, the Gibbs phenomenon error, which arises from finite resolution and is associated with convection across a grid, is going to be present regardless of the solution algorithm. This Gibbs error is also not Galilean invariant and is a function of the representation and resolution, not the solution algorithm. In fact, the non-Galilean feature of monotone algorithms is designed to cancel this non-Galilean aspect of the solution arising from the Gibbs phenomenon. The composite interaction of a monotone algorithm with the representation gives a solution which is essentially Galilean invariant.

Conversely, any algorithm which itself is Galilean invariant will be unable to cancel the Gibbs error without extensive diffusion. Thus the resulting solution will either be highly diffusive or will not be even approximately Galilean invariant. In DNS applications the real viscosity provides adequate diffusion for the resolved velocity field. Here, adequate diffusion is defined to be at least as much as occurs in first order upwind algorithms. The price of this approximate Galilean invariance, which equates closely with physical monotonicity, is a severe limit on the Reynolds number which can be reached. In multimaterial flows or flows with contact surfaces, viscosity alone is not generally adequate to ensure Galilean invariance for physical variables other than velocity which is smoothed by viscosity.

In the remainder of this section we consider how monotone algorithms function as an integrated subgrid model and do this in the context of the five desired properties of a subgrid model identified in Section 2 above. Properties P1 (generality) and P2 (conservation)
are built into the formulation of the basic flux-corrected convection algorithm. Property P3 (minimal contamination of resolvable scales) is the ensatz underlying monotone convection algorithms. Sharp gradients in fluid profiles are convected with minimal numerical smoothing consistent with keeping positive quantities positive and keeping physically monotone profiles monotonic to the maximum numerical resolution allowed by the grid. Since minimal dissipation is used to do this and since this flux-correction or flux-limiting is highly localized, it follows that monotone algorithms also entail minimal contamination of the well-resolved long wavelengths. As noted above, detailed measurement [41,45] gives an effective dissipation scaling roughly as the fourth power of the spatial scale. This means that flow structures that are 10 times larger than the grid scale structures, which must be dissipated strongly, feel 100 times less residual dissipation than the same macroscopic structures would be subjected to in a conventional LES or DNS model where the short scales are controlled by a viscous or eddy diffusivity term in the equations. Thus one can expect these large structures to be accurately convected for Reynolds numbers up to 100 times larger.

Existing subgrid models, until recently, have been generally limited to constant density, incompressible, non-reacting flows on uniformly spaced meshes, effectively violating property P1. Current developments of LES for compressible and reactive flows, even with Favre averaging and filtering, have many more nonlinear closure terms to deal with and correspondingly more phenomenologies with free constants to be calibrated. These difficulties, at least to first order, do not plague FCT and other MILES approaches.

Property P4 (physical subgrid mixing) is enforced by FCT through the residual local dissipation left to enforce property P3. This feedback clearly loses information about the unresolved small scales but other subgrid models also lose this information as they generally lack the random feedback effects. Furthermore, FCT and other MILES algorithms do allow deterministic backscatter from the short wavelengths to the resolved scales through the intermittent and localized nature of the flux-correction. Leith [19] has proposed and considered models for stochastic backscatter which also point the way toward their inclusion in MILES algorithms. A random local excitation of long wavelengths is required and could easily be added, for example, since the exact amount of FCT flux-correction is known at each cell interface, but this has not been done to date.

Property P5, that the subgrid model match smoothly onto the LES model, is perhaps
the most attractive and compelling aspect of using monotone convection methods for LES. A consistent and integrated viewpoint is used to convert unresolvable fluid dynamics (and grid resolution limitations) into the subgrid fluid dynamic fields. Even with spatially and temporally varying grids, a severe problem for conventional LES, the residual long wavelength transport from the MILES flux limiters acting on the cascading subgrid-scale fluctuations is included causally and consistently in the composite model while ensuring Properties 1–4.

The intrinsic filter in monotone algorithms is problem and grid dependent but, with increasing resolution, the numerical solution converges to the solution of the underlying partial differential equations being solved [46]. This means that the well resolved field solutions may differ at most slightly from the exact (laminar) solutions of the equation set being modeled. Inverting this built-in filter (i.e., defiltering) is not possible except statistically but this inversion also should not be necessary for quantities which depend only on the well resolved scales of motion. Conversely, the FCT filter can be applied to experimental data or to theoretical models if more detailed comparison with computations of the resolved scales is desired and the only information available depends significantly on the unresolved scales.

In the next section we discuss series of simulations performed specifically to look at the convergence of FCT with increasing resolution, to look for the minimum of entrainment expected at an intermediate resolution, and to determine how much added eddy transport may be needed or desirable.
5. Calibrating Flux-Corrected Transport for Large Eddy Simulation

A series of three-dimensional simulations have been carried out to study the convergence of a MILES algorithm, FCT, as resolution is increased for a high-$Re$ flow. These simulations expose the action of the built-in subgrid model as more and more short-wavelength structure is resolved in the solution. The problem chosen is the turbulent entrainment of quiescent background air into a fast but subsonic cylindrical jet. This system highlights the intrinsically three-dimensional phenomena which contribute to enhanced mixing, treats a spatially inhomogeneous problem which is of interest beyond its implications for LES, and prepares the way for reactive jet and detonation simulations using efficient models of exothermic combustion chemistry. Turbulence is often localized spatially, working its way into neighboring regions of potential flow through a fairly sharp time-evolving interface where the vorticity drops rapidly to zero. A number of the systems described in Section 4 have this property as does the fast but subsonic jet simulated here.

In this problem the jet is composed of air at standard atmospheric temperature and pressure, the same conditions as in the background. The gas constant $\gamma = 1.40$ and the jet centerline velocity is $V_{jet} = 150$ m/s, giving an initial Mach number $M = 0.452$. The jet velocity is constant inside 1.0 cm radius, decreases linearly with radius from $V_{jet}$ to zero between 1.0 cm and 1.4 cm, and is zero outside $r = 1.4$ cm in the undisturbed background gas. Thus the initial vorticity thickness is $\delta = 0.4$ cm and the initial jet radius is $R_{jet} = 1.2$ cm. This shear layer is resolved with at least two or three cells in the coarsest-resolution grid and with eight to ten cells in the fine-grid cases. This ensures that the most unstable modes [42,43] are well resolved by 20–40 cells per wavelength and thus prevents nonlinear saturation of different, progressively higher-frequency modes from dominating the evolution of the system as numerical resolution is increased.

The system considered here is periodic in the $X$ (streamwise) direction to maximize the use of grid resolution and to simplify the computations for the TMC Connection Machine. A number of simulations were carried out with a system periodicity length of $L = 12.8$ cm. Most of the simulations started with a relatively large amplitude mode 3 helical perturbation of the circular jet. Spatial inhomogeneity enters this problem through
the boundary conditions as the grid is stretched away from the vicinity of the jet in the Y and Z directions. Different transverse gridding, different formulations and amplitudes for the initial perturbations, and different ways of evaluating the macroscopic entrainment and other diagnostics were formulated and tested. The domain was structured so that the jet flows within a central core of uniform cubical cells.

Physically identical jets were simulated using the Naval Research Laboratory Connection Machine (CM) with different computational meshes to test convergence as more spatial scales are resolved. The first grid has $128 \times 64 \times 64$ cells of size 0.1 cm which will be referred to as medium resolution. The second grid has $256 \times 128 \times 128$ cells of size 0.05 cm, referred to as fine resolution. To allow longer calculations without degradation from stretched cell regions, several tests were performed to pick a suitable gridding structure. The computations reported here were run with stretched cells six times larger at the side boundaries and with uniform cells occupying the central 70% of the grid in the Y and Z directions. Thus the system width is about 12 cm but the jet moves some distance sideways before encountering stretched cells. As the jet becomes highly convoluted in this grid, however, fluid eventually moves into the stretched cell region and the computation then loses resolution.

A version of the code was developed to run simulations on a Convex C2 with a $64 \times 32 \times 32$ mesh of 0.2 cm cells, called coarse resolution. Gridding and initial conditions are identical with the CM calculations but the Convex was used because the $64 \times 32$ cross sections of the coarse grid are too small to make effective use of the CM.

Table 1 shows the resolution parameters and scaling ratios for these simulations of MILES convergence and for some related, finer resolution grids which can eventually be used to extend the results here to even higher resolution. For the simulations reported here the jet was perturbed in a mode-3 helical pattern at a wavelength of 4.27 cm which is $\lambda = 3.56 \times R_{jet}$. This perturbation rotates three times in traversing the length of the system and was originally implemented by displacing the column helically off axis about 0.05 cm with zero transverse velocity. This way of initiating the system was used only for the medium-grid mode amplitude results shown in Figure 2 and discussed below. For the
series of three MILES runs reported below testing convergence, the displacement was set to zero but the jet was given a velocity through the background transverse to its axis with a fixed magnitude but with a direction rotating three times in traversing the system. The transverse velocity in each cross section had a constant transverse velocity for $r \leq R_{jet}$ inside $R_{jet} = 1.2$ cm and an incompressible (dipole) recirculation in the background gas outside the jet. In both initialization procedures the perturbation level corresponded to a couple of percent of $V_{jet}$ while the density and pressure were left unperturbed.

Neither of these two ways of perturbing the straight jet column is an exact eigenmode of the linear system so some ambiguity exists as to how to determine the evolving mode amplitude. Mode amplitudes are approximated here by averaging the transverse displacement of the fluid inside of $R_{jet}$ at each axial cross section of the computational domain and then by Fourier analyzing these rows of average $Y$ and $Z$ transverse velocities as a function of $X$. This Galerkin-like procedure is most meaningful in the linear regime but simplifies the question of how to analyze an inhomogeneous three-dimensional field and bypasses the lack of accurate eigenfunctions for this particular problem.

The third mode of the system, $\lambda = L/3$, was chosen so that physically identical flows could be initialized with different cell sizes, as indicated by Grids 2, 4, 6, and 8 in Table 1, while catering to the CM's preference for grids where the number of cells is a power of two. By increasing the perturbation wavelength by 50% and reducing the cell sizes a corresponding amount, the chosen wavelength could be maintained by switching to mode 2 of the system. Further, the increase of a factor of two in the number of cells perpendicular to the jet axis accommodates the better resolved flow entirely within uniformly spaced cells. Table 1 presents the resolution parameters of a sequence of cylindrical-jet runs testing LES convergence by comparing entrainment time histories. Simulations for Grids 1, 3, 5, and 6 were performed for this paper. Grids 6, 7, and 8 have been made possible by memory upgrades to the CM and will be used for revised MILES calibration runs in the future.

Using the third mode of the system as the primary perturbation allows modes 1 and 2 to appear later in the simulations as subharmonics arising in large-scale vortex merging, thus bringing longer as well as shorter wavelengths into the flow. Since, however, the
amplitude of these subharmonic modes was initially set to zero in the simulations presented here, these modes began to grow significantly from roundoff errors only toward the end of these simulations. These subharmonics have yet to be explored systematically though several short auxiliary runs were carried out at medium resolution (Grid 3) with modes 3, 4, and 5 perturbed simultaneously while testing the gridding.

Since the primary goal of these simulations is the turbulent mixing which follows the nonlinear saturation of the initial Kelvin-Helmholtz instability, the linear phase of the problem is treated more crudely than if linear theory and the linear behavior of the code were being compared. Large initial perturbations are used to speed the transition to the important nonlinear mixing regime between 1.0 ms and 2.5 ms (e.g., Figures 2 and 6). The initial perturbation was not a perfect eigenmode and the nonlinear limiter in FCT, also used during the linear growth phase, has the effect of generating the higher harmonics of the initial shear layer perturbation with the same symmetry as mode 3, i.e., modes 9, 15, 21, etc., though these modes were not present in the initial perturbation.

The amplitudes of these shorter-wavelength modes as well as of mode 3 are shown as a function of time in Figure 2 for the early Grid-3 simulation (dashed curves) initialized with a transverse helical jet displacement and for the later Grid-5 run (solid curves) initialized with a helical transverse velocity perturbation. The mode-9 shorter wavelength instability was about 14 cells long in the medium resolution grid and 28 cells long in the fine grid. It appears at low amplitude initially but has a larger linear growth rate than the primary, mode 3. Flow visualizations during the linear growth phase make it clear that mode 15 (and also modes 21, 27, etc., not shown in Figure 2) are all phased to mode 9. An example of such a visualization is shown in Figure 3 for a higher-resolution run using Grid 6. The figure shows grey-level contours of the three velocity components on a cross section through the center of a higher-resolution (256 x 256 x 256) jet near the end of the linear growth phase. In this figure, the fully-developed Kelvin-Helmholtz vortices of shorter-wavelength can be seen clearly superimposed on one wavelength of the primary perturbation which is still growing. In the simulations of Figure 2, mode 9 saturates at considerably lower level than the primary mode 3 and certainly is superimposed on top of the primary. These shorter-wavelength modes constitute the secondary instabilities of turbulent cascade and
contribute significantly to the additional entrainment in the fine-grid FCT run relative to the medium-grid run when the flow is fully developed.

Figure 2 shows an identifiable period of linear growth of the primary mode 3 in both medium resolution (Grid 3) and fine resolution (Grid 5). The medium-grid run, initialized with a jet displacement (but zero transverse velocity), has both growing and decaying eigenmodes present initially with the same amplitude. This explains the zero slope of the curve (dashed with boxes in the figure) at $t = 0.0$ s as the composite mode amplitude departs quadratically from its initial value. Changing the form of the helical perturbation from spatial displacement to transverse velocity, as described above, had the effect of reducing the initial mode amplitude perturbation by almost a factor of two. The initial pressure perturbation is still zero in the fine-grid case, so both growing and decaying modes are still present with comparable amplitude. Therefore, linear growth is only seen after one or two e-folds ($\tau = 0.17$ ms) have elapsed. Nevertheless, more than an order of magnitude of linear grown is seen, as emphasized by the straight line. Furthermore both resolutions clearly show the same linear growth of mode 3 despite their differences in resolution and method of initial perturbation. This figure also shows clearly that the effective subgrid model intrinsic to FCT is not damping the growth of either mode 3 or mode 9 appreciably.

Comparison of the measured growth rates with published results from linear theory is complicated because of the differences between this physical problem and those found in the literature. Martin and Meiberg [92] simulated the temporal evolution of helically-perturbed jets on the linear and early nonlinear regimes using a vortex-dynamics method but do not present a linear growth-rate analysis. Michalke and Hermann [42] and Michalke[43] have performed linear stability analysis for axisymmetric jets. None of the shear-layer profiles considered by these authors matches the one used here. Their data does not include growth-rate curves for $\frac{R_{\text{ext}}}{\delta_s} \approx 18$ ($\theta_s$=initial momentum thickness) which characterizes the three simulations in the convergence series performed for Figures 2, 6, and 7. Qualitative comparison is possible, however. The spread of linear growth times, estimated from $\frac{R_{\text{ext}}}{\delta_s} \approx 20$ results, is $0.09 - 0.15$ ms for mode 3 and $0.05 - 0.10$ ms for mode 9. These numbers are consistent with the longer growth time in the simulations, $\tau = 0.17$ ms. Further, theory says mode 6–8 should be fastest growing but the difference in growth rate between
modes 3 and 9 would only be about 50%. More detailed investigation of the linear fidelity of these simulations is unwarranted without better results for comparison and simulations performed with a considerably smaller initial perturbation.

Nevertheless, the primary perturbation clearly grows linearly at first. During this phase, very small nonlinear effects from the flux limiter trigger shorter wavelengths which grow faster. The resulting saturation of the Kelvin-Helmholtz instability rapidly becomes extremely complex (turbulent). Figure 2 shows an apparent delay of 0.1 ms in the growth and saturation of mode 3 in the fine-grid case relative to the medium-grid case. This occurred because the transverse initial velocity of 100 cm/s (solid lines in Figure 2) corresponded to about a two times smaller displacement of the jet in terms of the initial perturbation amplitude than used for the medium grid (dashed lines). It is also clear that the mode 9 secondary component of the unstable shear layer saturates at a higher level and at a later time in the fine-grid case.

These differences are explained by the changed resolution. In the fine grid, the effects of the nonlinear flux limiter are less significant than in the medium or coarse grids and thus the secondary mode amplitudes are initially smaller. They eventually grow to a higher level in the fine-grid case however, because there is less nonlinear damping of the short wavelengths by the built-in subgrid properties of the FCT. This can be seen at about 1.5 ms where the fine grid mode 9 peaks at a level of about 300, twice as high as the medium-grid case and a full 0.5 ms later. Since energy is conserved in these flows, the additional energy in the fine grid mode 9 has to come from somewhere; a corresponding small reduction in the mode 3 amplitude is seen at about the same time, relative to the medium-grid case. This is significant because a macroscopic quantity such as entrainment, which depends on both modes 3 and 9, may actually increase more slowly at first on the fine grid relative to medium and coarse grids due to the necessity of populating all resolvable scales of the turbulence from a fixed amount of available kinetic energy.

Figure 4 shows several cross-sections from the medium-grid solution (top two panels) and from the fine-grid solution (lower panel). Grey-level contours of the jet fluid are shown as it mixes with the background air. The effects of mode 9 are superimposed on mode 3.
here. Vorticity of one sign rather quickly appears on the other side of the jet as it is pulled around by the nonlinear helical structure. Thus the nonlinear gradients are steepened further by the stretching inherent in this flow, and this leads to additional secondary and tertiary small-scale structure. At medium resolution, some of this fine-scale structure can be seen but the fine-resolution calculation shows considerably more structure. Viewed globally, the nonlinear Kelvin-Helmholtz instability in this geometry seems to deform the jet into a roughly helically symmetric core with a thin, helical shroud nearly encircling it. This shroud can be seen as the thin strips of fluid indicated in all cross-sections in the lower panel of Figure 4. Between the shroud and the core is a layer of engulfed fluid which is essentially vorticity free and which is entrained by the scales of turbulence resolved in these FCT simulations.

The shear-layer entrainment velocity gives the rate of propagation of the interface between rotational and irrotational fluid. On the smallest scales, viscosity acts to propagate vorticity into the irrotational fluid. However, the entrainment velocity is controlled by the speed at which the contorted interfaces of the largest scales move into the surrounding fluid [25]. More generally, the entrainment velocity gives information on the rate at which the free streams become mixed as they join the shear layer. Approaches to measuring fluid entrainment have typically been either based on the vortical content of the fluid [72] or obtained approximately by examining the spread of the velocity profile, in terms of volumetric fluxes [93], or evaluating the so-called passive scalar entrainment [94], which is closest to the approach used here.

Here the study of convergence with resolution is based on examining the temporal evolution of the jet volume, which is taken as a convenient macroscopic measure of entrainment. This entrainment volume was used to measure the size and cooling rate of channels generated by lasers, lightning bolts, and charged-particle beams in air [58,59,82,83] as described in Section 4. The sensitivity of the present convergence studies to this particular definition of entrainment, as well as its dependence on the choice of initial conditions, deserves further study and will be compared with other measures in investigations to be reported elsewhere.
A passive scalar $\phi$, initialized with the same linear profile ($0 \leq \phi \leq 1$) through the shear layer as the $X$-component of the velocity, is used to mark the fluid initially in the jet. The volume of fluid in the entire system that has $\phi \geq \epsilon$ defines the entrainment volume for a given $0 < \epsilon < 1$. It is calculated by summing up the volume of cells satisfying the criterion $\phi \geq \epsilon$. Initially this entrainment volume is approximately $\pi R_{jet}^2 L$. Figure 5 shows an example of this overall entrainment diagnostic using $\epsilon = 0.02$, 0.05, and 0.10 plotted as a function of time for a recent fine $256^3$ calculation (Grid 6). As the jet fluid spreads, the average value of the passive scalar in the jet plus mixed region drops. Initially $\epsilon = 0.5$ marks the interface between the jet and the background at $r = R_{jet}$. Lower values of $\epsilon$ give somewhat larger volumes. As mixing progresses, however, the entrainment volume quickly doubles, at which point a diagnostic with $\epsilon = 0.5$ would soon miss much of the mixed fluid because the background air would tend to dominate the mixture almost everywhere. A lower value of $\epsilon$, as used here, allows calculations to progress longer before the entrainment diagnostic loses its meaning due to dilution as the volume with $\phi \geq \epsilon$ begins to decrease.

The three levels shown in Figure 5 give almost the same entrainment volume initially but the small deviation grows as mixing creates a larger and larger volume of fluid with entrainment ratios between, for example, $\epsilon = 0.05$ and 0.1. The value $\epsilon = 0.05$ was used as a compromise for the cases here with attention paid to the bounding values of $\epsilon$ to ensure that the main body of the mixed fluid was not diluted out of the range of visibility. Figure 6 shows this entrainment-volume diagnostic computed as a function of time at the level $\epsilon = 0.05$ for the fine, medium, and coarse grid runs performed for this paper. The entrainment volume was normalized by dividing out the initial volume in each of the three runs because each of these initial volumes was slightly different. This occurs because the entrainment threshold $\epsilon$ falls across the profile of the passive scalar slightly differently on each of the three grids. Since a cell either does or does not satisfy the entrainment criterion, the corresponding sums are quantized and differ by about a percent. This became important because the entrainment, as can be seen, is so nearly equal in the fine and medium resolution cases.

With most linearly filtered LES algorithms, the correct entrainment for a turbulent
high-$R_e$ flow would be approached from above, as shown on Figure 1. The finite resolution solutions must be defiltered to convert a macroscopic quantity such as the entrainment to its infinite resolution value and that procedure is potentially unstable numerically. With MILES algorithms, the effective filtering is nonlinear and thus the nonphysical diffusion does not extend significantly to long wavelengths. The three curves in Figure 7, corresponding to 2.0, 2.2, and 2.4 ms in the evolution of the jet at the three resolutions chosen, demonstrate that the predicted minimum [3,4] is indeed found at intermediate resolution.

The minimum occurs because short wavelengths, which would provide some entrainment, cannot be resolved. The residual numerical diffusion present in high-order monotone algorithms is smaller than the eddy diffusivity of the turbulent flow, so increasing resolution increases the entrainment. The minimum is not very deep and is resolved with only three points in these simulations but it appears to be getting deeper as time progresses, another indication that the additional scales of vorticity in the fine-grid case are continuing to increase the entrainment-volume relative to the medium and coarse grid. The fact that the minimum is so shallow, at most 2–3% deep, is actually beneficial. A shallow minimum, while being correspondingly more difficult to measure accurately, means that virtually no additional subgrid-scale transport, as indicated schematically in Figure 1, is needed to get the “right” answer. The intrinsic subgrid model provided by FCT, at least for these cases, is very good.

6. A Hydrodynamic Analogy for Turbulence Modelling

In fluid systems, turbulence often begins as a large-scale response to unsteady external forcing or to macroscopic instabilities as they become nonlinear and restructure the otherwise laminar flow. Energy drives the complex flow first at long wavelength but then cascades to shorter and shorter wavelengths through a sequence of nonlinear couplings where it is eventually dissipated viscously. The flow of energy through the “inertial range” to the Kolmogorov scale [23–25] can be likened to the flow of water away from the center of a flat table on which it is being poured. This analogy provides a way to understand and visualize this turbulent cascade in the context of the different approaches to LES considered above. The three panels of Figure 8 depict this analogy schematically.
Driving a turbulent flow at resolved macroscopic wavelengths is like pouring water onto the center of a large flat table. The water flows radially outward, getting thinner and moving faster so the mass flow past any radius is constant. Increasing radius away from the center of the table is analogous to increasing wavenumber of the eddies in a turbulent cascade. The decreasing depth of the water is analogous to the decreasing energy content in each wavelength scale of the turbulence. The inertial range of the turbulent cascade is represented by the region between the vertical dashed lines where the profile is smoothly decreasing in Figure 8.a. The radius of the table and how the water eventually falls off the table, analogous to the viscous dissipation of turbulent energy at the Kolmogorov scale in very high-\(R_e\) flows, clearly does not significantly affect the depth of the water near the center of the table. In this hydrodynamic analogy, different possible contours at the edge of the table correspond to the different properties of various high-\(R_e\) Navier-Stokes models, conventional filtered LES models and MILES models.

In MILES models based on monotone convection algorithms, the nonlinear flux limiter acts is a built-in subgrid model coupled intrinsically to the short wavelength errors in the solution. Turbulent energy reaching the grid scale is extracted from the calculation and converted to the correct conserved quantities. This has the effect of curving the table edge sharply downward, as illustrated in Figure 8.b, so that the water can flow smoothly off at a finite radius without significant perturbations reaching the center of the table. The dissipation in MILES algorithms is physically matched to the grid-scale errors to minimize effects on long wavelengths which are accurately resolved.

With conventional, high-order CFD algorithms which are not monotone, dissipation is added through the physical viscosity. Thus a blocking or damming up phenomenon [13] occurs for high-\(R_e\) flows where the grid-scale fluctuations build up to nonphysical levels because the excitation cannot be removed at the rate it is generated. In this water-spill analogy, this energy-blocking effect is like a raised rim around the edge of the table at a radius corresponding to the grid scale of the computation. A layer of stagnant water would form out to the table edge below the height of this rim. To prevent this, conventional LES algorithms use the grid filter, often in the form of a spectral cutoff, to add appreciable smoothing to the equations. In addition, the eddy viscosity used to model the effects of
the subgrid scale turbulent stress, incorporates significant additional dissipation. In some models this added dissipation is sufficient to stabilize laminar flows which otherwise would be unstable as noted in [20,21]. The effect of filtering the equations or including significant physical or eddy viscosity is analogous to curving the surface of the table downward so the rim is now at the level of the table center.

7. Summary and Discussion

Above, it is argued that nonlinear monotone methods really have a built-in LES filter and a matched subgrid model which do minimal damage to the longer wavelengths while still incorporating, at least qualitatively, most of the local and global effects of the unresolved turbulence expected of LES. When properly formulated, a wide variety of these monotone convection algorithms transform unresolvable structure in the fluid field variables into the appropriate modifications of the resolved fields. This structure and the corresponding flow energy are pushed to short wavelengths by nonlinear convective effects and instabilities in the flow. Because global conservation is enforced through purely local algorithms, the grid scale variability is locally converted to the correct macroscopic variable averages. For example, kinetic energy entering the subgrid scales from the resolved motions is damped as the velocity fluctuations are dissipated by the flux-limiter in FCT. Since energy is conserved while kinetic energy decreases locally, the pressure goes up accordingly just as if physical viscosity at unresolved scales had converted the Kolmogorov-scale structures to heat. Of course the details (and short time delays) associated with the cascade through the unresolved inertial range is lost but this is accepted in all LES modelling. Indeed, this appears to be an area where fluid dynamics has been kind to us in the sense that the large scales do not appear to be particularly sensitive to the components of the flow which cannot be simulated.

Furthermore, these methods are quite capable of capturing quantitatively how much unresolved structure from the long wavelengths is actually present. Diffusion of the eddy transport type is automatically left in the flow, but the fluctuating, driving effects of random-phase, unresolved eddies on the large scales is missing unless specifically included as a subgrid phenomenology. A factor of two increase in the spatial resolution of LES and
MILES models will most likely bring more improvement in the accuracy of the well resolved scales than all the work in the world on the subgrid model of a more coarsely resolved LES model with the usual filtering procedure that contaminates the long wavelengths. Satisfying proofs of these statements have not been provided here, but work is underway to do exactly this.

Even with careful attention to the initial conditions, calibration of the long-term behavior of LES models is difficult because fine, medium, and coarse resolution calculations of what should be a single physical flow will deviate from each other rapidly due to details of the flow structure which is resolved in one case and not resolved in others. Although the macroscopic properties and averages of the flow may be perfectly deduced from a relatively coarse LES, it turns out to be nearly impossible to prove this because statistical averages involving only a few percent of the flow energy are generally being sought. Thus, comparison simulations must be run for a long enough time in a large enough volume on a statistically stationary problem to ensure that averages over the unavoidable but generally unimportant phase separations that appear in different cases will have smaller errors than the phenomena being studied. This usually means that many characteristic times of the largest scales of the system must be included in the average before statistical errors arising from this intrinsic variability are smaller than the quantities of interest.

With the recent addition of memory to NRL’s 16,000-processor CM and various other system upgrades, it is now possible to perform calculations with Grids 6 and 7 in the table, namely computations on the fine 256 × 256 × 256 grid and on an extra fine 512 × 256 × 256 grid. With a full 64K-processor CM, a grid with 512^3 cells is possible. The computations for Grid 5 took 18 seconds per timestep on NRL’s 16K processor CM using a specially coded version of the LCPFCT routines provided by R. Whaley [95,96]. Recent system improvements have obviated some of this optimization so 256^3 calculations now take about 80 seconds per timestep. On a full CM, calculation on a 512^3 mesh (Grid 8) would take about 150 seconds per step or about 24 steps per hour when integrating five fluid dynamic variables and one extra passive scalar in fully compressible gas dynamics. Full runs of 10,000 steps would therefore take more than 40 hours.
With additional diagnostics and modified initial conditions, these higher-resolution grids will be used to continue these investigations of the built-in MILES subgrid behavior. From the studies reported here and various additional tests performed along the way it is clear that problems must be chosen where the long wavelengths are continually pumped to provide a statistically steady base line without the intrinsic decay built into tests such as reported here and elsewhere. The problems chosen should also make better use of the available resolution than was done here because of the wasted cells at the edge of the system and the three essentially identical, replications of the system in the streamwise direction.

Using a more nearly homogeneous problem will also allow the meaningful use of Fourier transforms to determine spectra, where specific information about the missing short wavelengths can be measured rather than deduced as in the present paper. In addition, diagnostics based on the vorticity and on the volume of mixed fluid alone should be used to augment the information obtained using the volumetric entrainment as defined here. For our FCT algorithms, the time and space distribution of the unused fluxes of mass, momentum, and energy should be studied as measures of the unresolved subgrid dynamics. Differential measures of small scale dynamics should also be made repeatedly by starting medium and coarse grid simulations from the instantaneous state of a fine grid calculation and then differencing the results on the coarse grid after a short period of time. This approach would remove the effects of accumulated phase differences between the different resolution calculations at large scales from masking the effects of progressively better resolved small-scale dynamics.

Another variation would be to restart a fine-grid MILES run adding a local, relatively small-scale perturbation to the flow with a known energy and spectral content. This turbulent patch could then be followed for a short time by differencing the computation with the original run where the the patch was not added. By diagnosing the action of the flux-limiter, the cascade of the added energy off the grid could be monitored in time and space for comparison with expectations based on the dissipation rate of the associated Kolmogorov spectrum.
Since MILES methods have not usually been thought of as LES models, there are areas where extensive interpretation and verification are needed. It has to be demonstrated that the residual average transport at long wavelength from unresolved subgrid turbulence is large enough. Almost certainly additional eddy viscosity must be added to the minimal amount provided by the algorithms. The essentially random and fluctuating components of the subgrid fields are also missing from these integrated LES models as well as from other LES models, and should be included. The cell-averaged source terms which drive these fluctuations are available, however, as they are contained in the components of the fluxes removed in the nonlinear limiting process. Physical assumptions about the short timescale temporal behavior (cascade) and spatial characteristics of the unresolved motions have to be made. All that is known about the subgrid fields during the simulation has to be inferred from the resolved fields.

Source terms in the LES equations can be included for these subgrid fluctuations once what is scientifically appropriate has been decided. Research on the subject has been initiated by Leith [19]. This will be a phenomenological model but the goal of this approach is to require as little as possible in the way of subgrid terms be added to the underlying monotone algorithm. It appears that such terms should have a local and random aspect on the macroscale so that resolvable-scale flow instabilities and hence turbulent structure can be triggered by dynamics on the unresolved scales. This, also, is easy to do. Different subgrid augmentation algorithms should be tried for FCT once a baseline convergence behavior corresponding to Figures 1 or 7 is obtained. The goal of these augmentations would be to add conservative fluctuations of a random nature to the resolved-field calculations which are based on the unused fluxes identified by the nonlinear FCT flux-limiter. These augmentations would presumably be able reduce or eliminate the minimum in the entrainment vs resolution curve, making macroscopically correct simulations possible at even coarser resolution. Furthermore, these fluctuating subgrid source terms automatically will lead to additional macroscopic transport because the monotone flux-limiters will work on these subgrid-determined effects as well as on the macroscopic effects. In the simulations reported, a stochastic backscatter model, the subject of recent research elsewhere [19,70], was not included to augment the minimal
eddy viscosity provided by the monotonicity-preserving flux limiter built into the FCT algorithm.

Finally, to help understand these LES concepts and the comparisons of different approaches, an analogy with the flow which results from pouring water onto the center of a flat table was introduced. While this analogy is certainly not rigorous in any way, it makes cascade and eventual dissipation of turbulent energy at short wavelength subject to intuition and visualization.

Acknowledgements

Research on different fluid systems using this MILES, or under-resolved Navier-Stokes approach, has been supported by a number of sponsors including ONR, the NRL, DNA, DARPA and DOE. The authors also wish to acknowledge the input and inspiration received during many discussions with K. Kuwahara, P. Woodward, R. Dahlburg, J. Gardner, and K. Kailasanath over the last few years as well as the varied and creative contributions of colleagues at NRL and around the world to understanding complex and turbulent fluid phenomena through numerical simulation using monotone methods.
References


46. T. Ikeda and T. Nakagawa, On the SHASTA FCT Algorithm for the Equation \( \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\nu(\rho) \rho) = 0 \), *Mathematics of Computation* 33 (148): 1157-1169.


Table 1. Sequence of Cylindrical Jet Entrainment Runs Testing LES Convergence. Simulations using grids 1, 3, 5, and 6 were performed for this paper. Though not used for this paper, the data for grids 2 and 4 are included to show how intermediate resolution results could be determined. Grids 7 and 8, which would give an additional factor of two or three increase in resolution beyond the runs reported, are now possible due to memory upgrades to the Connection Machine.

<table>
<thead>
<tr>
<th>Case</th>
<th>Configuration</th>
<th>$\delta x$ (cm)</th>
<th>mode</th>
<th>$h/\delta x$</th>
<th>$R_{jet}/\delta x$</th>
<th>$\lambda/\delta x$</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$64 \times 32 \times 32$</td>
<td>0.20</td>
<td>3</td>
<td>2.0</td>
<td>6.0</td>
<td>21.33</td>
<td>coarse grid, Convex C2 velocity perturbation</td>
</tr>
<tr>
<td>2.</td>
<td>$64 \times 64 \times 64$</td>
<td>0.15</td>
<td>2</td>
<td>2.67</td>
<td>8.0</td>
<td>28.44</td>
<td>coarse grid Convex C2 various tests not reported-</td>
</tr>
<tr>
<td>3.</td>
<td>$128 \times 64 \times 64$</td>
<td>0.10</td>
<td>3</td>
<td>4.0</td>
<td>12.0</td>
<td>42.67</td>
<td>medium grid, 16K CM-2 perturbed in both position and velocity</td>
</tr>
<tr>
<td>4.</td>
<td>$128 \times 128 \times 128$</td>
<td>0.075</td>
<td>2</td>
<td>5.33</td>
<td>16.0</td>
<td>56.89</td>
<td>medium grid 16K CM-2 various tests not reported-</td>
</tr>
<tr>
<td>5.</td>
<td>$256 \times 128 \times 128$</td>
<td>0.05</td>
<td>3</td>
<td>8.0</td>
<td>24.0</td>
<td>85.33</td>
<td>fine grid, 16K CM-2 velocity perturbation</td>
</tr>
<tr>
<td>6.</td>
<td>$256 \times 256 \times 256$</td>
<td>0.0375</td>
<td>3</td>
<td>10.67</td>
<td>32.0</td>
<td>113.78</td>
<td>fine grid, 16K CM-2 velocity perturbation</td>
</tr>
<tr>
<td>7.</td>
<td>$512 \times 256 \times 256$</td>
<td>0.025</td>
<td>3</td>
<td>16.0</td>
<td>48.0</td>
<td>170.66</td>
<td>extra fine grid 32K CM-2 not yet run-</td>
</tr>
<tr>
<td>8.</td>
<td>$512 \times 512 \times 512$</td>
<td>0.01875</td>
<td>2</td>
<td>21.33</td>
<td>64.0</td>
<td>227.33</td>
<td>extra fine grid 64K CM-2 not yet run-</td>
</tr>
</tbody>
</table>
Figure 1. Convergence of LES with Increasing Resolution
Figure 2. Evolution of Helical Perturbation Mode Amplitudes.
Figure 3. Cross-Sections from a 256 x 256 x 256 Jet Computation Showing Saturated Short-Wavelength Modes Superimposed on the Primary Perturbation.
Figure 4. Visualization of Nonlinear Evolution of a Helically Perturbed Jet.
Figure 5. Entrainment in $256 \times 256 \times 256$ FCT Simulation.
Figure 6. Circular Jet Entrainment vs Time.
Figure 8. LES Water Spill Analogy.