Challenges for Multiscale Large-Eddy Simulation of Application Systems: Gas Turbine to Scramjet

S. Menon and S. Srinivasan
School of Aerospace Engineering
Georgia Institute of Technology, Atlanta, GA, USA
E-mail: suresh.menon@ae.gatech.edu

Abstract. Simulation of turbulent combustion in operational devices is challenging due to the wide range of scales involved: from the molecular scales where mixing and combustion occurs to the large-scales of the device that control the overall processes. Interactions occur over the entire range of spatial (and associated temporal) scales, and models optimized to work in one specific region of application invariably fail when another area is simulated. For design studies of the next-generation low-emission gas turbines, simulation tools must allow the user to systematically adjust the operating condition from stable combustion to unstable combustion, without requiring any ad hoc adjustment to the model. Fuel-air mixing, shock-turbulence-flame interactions, ignition delay and flame holding are of interest in scramjet combustion. Described here is a multiscale, two-phase simulation approach that is being used to investigate these applications within a single simulation framework. Also described are computational challenges involved in strong scaling in multicore systems, in parallel optimization of the complex-geometry multiblock, and in improving efficiency of data acquisition and processing.

1. Introduction
Desirable features for next generation gas turbine engines for civilian applications are combustion efficiency, compact combustor size, reduced emissions, and stable combustion in the lean limit. For military applications, next generation engines are expected to operate under very high pressures using supercritical fuel. Currently, improving liquid fuel atomization and increasing fuel-air mixing are considered key issues to achieve efficient combustion within a compact zone. But the structure of the complex three-dimensional, swirling fuel-air mixing layers is very difficult to resolve using current experimental and numerical methods. From the numerical standpoint, since fuel atomization and fuel-air mixing are both highly unsteady, conventional steady-state methods cannot be used to elucidate the finer details. And, although unsteady mixing process can be studied accurately by using direct numerical simulation (DNS), application of DNS to operational combustors is still untenable. A simulation approach that has the potential to deal with such design challenges is large-eddy simulations (LES), in which all scales larger than the grid resolution are numerically simulated using a space- and time-accurate scheme while the effect of scales below the grid resolution is modeled using a subgrid model. Although many LES studies have been reported, a general-purpose, validated LES approach for simulation of practical systems is yet to be demonstrated, although some of the methodologies discussed in this paper have shown a potential with some quantifiable successes.
With the advancement in computing power, the feasibility of using LES for time-accurate simulations of turbulent combustion for increasingly realistic geometries has continually improved over the past few years. Whereas in the past, 2D, axisymmetric, or simplistic 3D geometries have been used, forced by computational (turn-around time and resource availability) constraints, the advancement in computational hardware has empowered CFD modelers to minimize assumptions regarding geometry and move toward realistic configurations. The rapid advancement of hardware technology has, however, instigated a revisitation of models and their implementation in an effort to improve their accuracy and performance. For example, the paradigm shift toward a multicore architecture has motivated the reinspection of traditional parallelization strategies used in legacy codes. Similarly, model assumptions or even entire models whose use was motivated by computational factors face the prospect of being overhauled or potentially replaced by more physically robust variants considered expensive heretofore. In this paper, the challenges of performing LES in realistic geometries in a cost-effective manner are described, along with the possible implementation strategies developed in our group. As will be seen, the challenges span the phases of grid generation, preprocessing, postprocessing, and mathematical modeling, and no final solution still exists, leading to continuing reassessment of simulation strategies.

2. LES methodology

LES modeling for reacting flows requires two seemingly independent (but, as shown below, interdependent) model development aspects. By definition, filtering implies scales below the filter width are lost and must be modeled. However, closure requirements for momentum and scalar transport are contradictory, making closure problematic. In LES modeling of the momentum transport, the effect of the unresolved small scales (assumed to be mostly isotropic) on the resolved motion is modeled by using an eddy viscosity-based subgrid model. This approach is (somewhat) acceptable for momentum transport since (almost) all the energy containing scales are resolved and the unresolved small-scales (that are assumed to be isotropic) primarily provide dissipation for the energy transferred from the large scales. For reacting flows, however, molecular mixing and kinetics occur at the small unresolved scales, and thus modeling these interactions at the LES-resolved scales brings with it some additional challenges. Many strategies have been reported in the past and summarized in many publications (e.g., [1]).

The fully compressible, unsteady LES equations are used in all our simulations. The choice of compressible formulation, which includes the coupling of acoustic-vortex-flame interactions without requiring any model changes, is considered particularly relevant if LES is to be applied to practical combustion systems, such as a gas turbine combustor, internal combustion engine, and rocket motors, that are not only confined but operate at high pressures and in which predicting combustion instability and lean blowout are some of the fundamental challenges in LES predictions. The interaction between acoustic waves, vortex motion, and flames involves a wide range of time and length scales and this range depends not only on the actual source (i.e., compressibility, shear layer separation and roll-up or heat release) but also on the configurational geometry. In confined domains boundary reflections can introduce another variable into the interaction process [2], and thus geometrical features have to be included with reasonable fidelity in the simulation model.

The numerical model employs a finite-volume algorithm based on an implicit box filtering scheme [3]. The subgrid scale terms for the momentum equations are closed using the conventional eddy-viscosity hypothesis. The eddy viscosity is computed by solving a transport equation for the subgrid kinetic energy and using a localized dynamic approach (called LDKM, hereafter) that employs a simple scale similarity approach to determine the model constants (for momentum and energy transport) as opposed to the Germano’s identity [4, 5, 6]. As shown elsewhere, the former approach is robust and stable locally even in complex domains. The species transport equations, when solved in the filtered form, are closed using a gradient closure-based approach via an equivalent eddy diffusivity approach. However, when a subgrid simulation approach—the Linear Eddy Model
(LEM)—is used, these transport equations are solved in their unfiltered form in the subgrid (called LEMLES, hereafter) [1, 7]. While more equations are solved in the LEMLES approach per LES cell compared with the LES-based approach, this method is nevertheless efficiently scalable with CPUs and has been shown to be feasible for realistic applications [3]. In LEMLES, since the molecular and reactive processes are simulated directly within the subgrid, the issue of reaction rate closure is side-stepped as the laminar rates can be used directly in the subgrid. This scalar mixing/combustion model has been demonstrated to capture the extinction/reignition phenomenon and shown to work for a range of combustion regimes without requiring adjustment [8, 9, 10]. More recent work has developed an artificial neural network (ANN) technique based on parametrization of the full reaction space in terms of finite rates for each species [11]. Both laminar ANN for use within the LEM and a turbulent ANN (TANN) for the filtered reaction rates have been developed and applied to extinction-reignition in diffusion flames and compared with high-resolution DNS. Good agreement with data has been obtained [12].

The spatial discretization for the gas-phase is multiblock, structured, and finite-volume based, second-order accurate in time, and either second- or fourth-order accurate in space. For supersonic applications, or for problems with strong discontinuities, a hybrid approach combines locally and dynamically a MUSCL scheme with high-order reconstruction with the central type smooth flow solver. As shown earlier, both moving shocks and turbulent shear flow are captured accurately and without smearing [6, 5, 13]. For two-phase flows, the particles are included by using a Lagrangian tracking method, which is necessary since in most reacting two-phase systems the location of the spray and its vaporization, mixing, and reactions are critical to the design goals. However, the simulation of the liquid phase using full two-way coupling with the Eulerian gas phase solver via source terms for mass (vaporization), momentum (drag) and energy (heat and mass transfer) is particularly challenging for LES modeling since particles too are considered much smaller than the filter width and, therefore, some sort of LES modeling is needed. In our approach, we consider a baseline stochastic separated flow model for simulating particle-turbulence interaction, and primary and secondary breakup models are used to model the liquid flow in the near-injector field [3, 14]. Models for the dense regime and for collisions are also included. However, this area still requires significant research to develop generalized LES models. The particle equations for motion, size and temperature are time integrated using a fourth-order Runge-Kutta method. A parcel approach, wherein multiple particles with initially identical properties are tracked en masse for the rest of the simulations, is used to bring down the computational cost without losing accuracy [15].

Given this capability, we have been able to address a few challenging issues in combustor modeling, lean blowout, and combustion dynamics in complex configurations and in configurations/conditions of industrial interest. Here, we discuss some of the LES and modeling challenges that were addressed in these studies and also issues that are still unresolved.

3. Complex Geometry Modeling Challenges
3.1. Grid Generation
Complex geometries imply a proper grid is needed. In our approach we consider structured multiblock grids (with or without adaptive mesh refinement) as the viable approach opposed to the unstructured approach popular in commercial solvers. Both approaches have pros and cons, but for high accuracy in the combustion region we believe a structured grid offers an advantage at this time. However, this choice implies significant overhead in grid generation that involves two important steps: determining the block topology and gridding the individual blocks. Typically, the blocking process is influenced by the topology of the geometry. A realistic geometry comprises numerous but volumetrically small components such as injectors and swirl blades and fewer volumetrically large components such as the combustor, inlet plenum, and injector housing assembly. Single blocks are assigned to the minor geometric features, whereas the larger components are divided into multiple blocks so as to bring the total grid-point per block to an optimized number. In
addition, features such as recesses or cavities in the geometric domain may result in the splitting of blocks to eliminate unmodeled components (such as the solid interior of a bluff-body or a swirl blade). All these elements can be seen in an example of a GE CFM56 type dual-swirl combustor in Fig. 1(a) where the eight primary swirlers constitute individual blocks and similarly, the flow-region between the eight secondary swirlers. The combustor shown in Fig. 1(b) connected to this assembly is divided into a larger number of blocks using axial and radial cuts to optimize the grid density per block.

![Grid for the CFM56](image)

Figure 1. Grid for the CFM56

For these complex geometries, a block structured approach is used in which the mesh within each block is structured whereas the blocks themselves are oriented in an unstructured manner. The simplest example is the “O-grid” where the central block is surrounded by four instead of eight blocks. The O-grid is typically employed in blocking cylindrical geometries (injectors, combustor, etc.) because it eliminates the centerline singularity for a purely structured grid while also improving the overall quality of the grid.

Another block-structured topology of the 9-injector lean direct injection (LDI) combustor is shown in Fig. 2, which is currently being simulated to study cup-to-cup interactions. The overall grid is a result of merging individual cups represented by the grid in Fig. 2(a) with the resulting grid-section at the dump-plane as shown in Fig. 2(c). In geometries such as this, the requirement of wall resolution near the vanes determines the overall grid size in the azimuthal direction. Hence, the azimuthal grid density increases with the number of blades for a swirl combustor and since the grid is structured, the azimuthal resolution at the swirler section is the global azimuthal resolution as well. If the grid clustering and the O-grid blocking are not performed carefully, the global minimum time step restriction can originate from the Cartesian mesh of the central block instead of critical regions such as the shear-layer. In order to avoid such situations in some cases, a grid-coarsening technique in which the radial and azimuthal resolutions are progressively reduced with decreasing radius is adopted resulting in a grid with hanging nodes. Such a strategy is shown in Fig. 3 where a 2-to-1 coarsening ratio is employed in the radial and azimuthal directions to reduce the mesh size and increase the global minimum time-step. Detailed analysis shows that the high-order interpolation used here allows acoustic waves and shear flow to pass through the grid interface without any numerical noise. In this example, 1106 blocks were used for a total of 18 million grid points. The uncoarsened mesh, in contrast, would have required over 45 million grid points. Although the latter resolution may be reasonable if large systems are available, with the coarsened mesh, the cost of one flow-through time is still around 13,000 single-processor hours on a dual-core PC cluster (circa 2009) for single-phase flow. Note that typically, 6–10 flow-through times have to be simulated to obtain statistical data.
Figure 2. Block topology and grid for the 9-point LDI combustor

The nonuniform mesh of the above strategy introduces modeling issues related to LES filtering. For example, while the filtering of variables from a fine to coarse mesh is straightforward, the procedure applied in the reverse direction is fraught with uncertainties since this essentially necessitates “unfiltering” of variables before their interpolation to fine grid. The commutation error of filtering and spatial derivative and the errors resulting from interpolation and filtering operation are areas of active research, as are issues related to the implementation of the LDKM model, LEM and Lagrangian particle tracking across the fine/coarse interface. Note that these issues are also of relevance to LES using AMR.

Figure 3. The grid coarsening technique used to couple a swirler assembly and a combustor

3.2. Multiblock Parallelization

A domain decomposition strategy is used for parallelization in which regions of the computational domain are allotted to different processors in a load-optimized manner using the Message Passing Interface (MPI) paradigm. Our current strategy is to have a system-independent implementation to allow portability to all parallel systems (e.g., IBM, Cray, SGI, PC clusters) rather than optimize for a single architecture. This has resulted in some compromises but allowed flexibility in simulations using in-house PC clusters as well as petascale systems when availability arise.
For multiblock geometries, the space decomposition is implicitly performed during the grid-generation stage when the block topology is established by the gridder. Given the large number of blocks for realistic geometries, it is not feasible to perform simulations using codes based on a single block per processor framework. In order to address these concerns, the code’s framework is implemented in a manner such that the computations can be performed in a block-specific manner with memory allocations for flow-variables specific to blocks. This framework allows multiple blocks to be accommodated in the same processor.

Communication in multiblock codes entails intraprocessor communications as well as MPI-based interprocessor communications depending on whether a block neighbor resides on the same or a different processor. Load balancing is then determined by striking a compromise between the inter-processor communication load and the cumulative communication load of blocks within the processor. Here, the grid count of the most dense block determines the limiting speed at which a domain-decomposed code will execute (i.e., at the speed of the most dense block). In geometrically simple configurations (such as those involving just the combustor), nominally uniform load-balanced blocking topologies are possible. For complicated configurations, however, the preponderance of small blocks makes load balancing a difficult task. One method to load balance a multiblock structured grid is to use a multi-constraint partitioning scheme. The two objectives are achieving optimal load balance and minimizing communication between processors. This can be accomplished with the METIS software suite [16]. A graph representing the grid is created with node weights representing the number of computational operations for each block and edge weights representing communication load between blocks. The relative weights between nodes and edges must be determined based on the both the computer hardware and algorithms being used. The METIS software is then utilized to complete the multiconstraint partitioning. The success of this method lies in the accurate determination of node/edge weighting and the size and complexity of the grid, as METIS works better for larger more complicated problems but produces poor results on relatively simple problems.

![Figure 4. Scalability of the solver](image)

The parallelization procedure for LEM is consistent with the above-described framework.
in the sense that the LEM parallelization strategy is also domain-decomposition based. The communications arise from the splicing procedure, which models the large-scale transport of mass and energy between LES cells. Since the subgrid structure is preserved during splicing, entire subgrid fields of the neighboring cells are transported and temporarily staged in the ghost cell locations of the LES cell (i.e., 6 subgrid fields in the 6 LES ghost cells). Based on the magnitude of the influxes on each cell face, whole or partial LEM cells are debited from the subgrid field of the appropriate ghost cell corresponding to the flux direction and credited to the host cell’s subgrid field (close to the beginning of the 1D field to obey a first-in, first-out convection procedure within the LES cell). The outfluxing operation is performed in an analogous manner [7].

Since the subgrid is resolved in LEM via a 1-D line where the scalar equations are solved, computational costs are about three to four times higher (for an LEM resolution of 24 cells per LES cell) than the cost of a conventional LES where only the resolved species equations are solved. For complex geometries however, two strategies can be explored. In the first, the choice of subgrid model (LEM or conventional LES) is made block-specific so that only those blocks in the critical regions of mixing and combustion are solved for using LEM. The conventional species transport equation with gradient closure approximation is used for the blocks in other noncritical regions. In this hybrid approach, load balancing takes into account the computational load imposed by these LEM blocks. In an alternative approach, resolution of the subgrid is assumed to be variable, depending on the local turbulence intensities. In this manner, the overall calculation load of each block can be considerably reduced, since the global subgrid resolution is no longer based on the smallest subgrid-size.

The parallelization performance of a kernel code similar to the complete multiblock solver is shown in Fig. 4(a) on the IBM Blue Gene/P machine and on various multicore PC systems for a benchmark nonreacting DNS problem solved using a fourth-order finite-volume scheme and for a single species. The metric used here is the “speed-up” defined as the time per iteration on one processor divided by time per iteration on \( N \) processors and the linear scaling for a range of grid/core densities spanning 1–4,096 processors indicates close to ideal performance of the code (access to systems larger was unavailable at that time). The code scalability on different multicore (dual, quad) architectures as shown in Fig. 4(b) confirm its scalability. Similar scalability has been obtained on other OEM systems such as the Cray XT-5, Cray XE6, and IBM Power6, showing consistent performance over a wide range of architectures. We note that no specific machine or compiler dependent optimization is carried out to ensure wider portability.

3.3. Particle Phase Parallelization

Two parallelization frameworks based on point-to-point or collective communications are possible for tracking particles or droplets, each with its own pros and cons. Both methods are based on domain decomposition wherein only particles residing in a certain processor, based on its spatial location, are processed by that processor. When particles cross the spatial boundaries of the host processor, communications are performed to transport them to their new host processors. It is here that the two communication frameworks differ. In the collective communication method, a global buffer is maintained with the array size proportional to the number of particles, and every processor retains a copy of this buffer while operating on only specific segments depending on the spatial location of the particles. For efficiency, particles belonging to a specific processor are stored in contiguous locations. At the end of each gas-phase iteration, the particles are communicated to the master-processor wherein book-keeping operations such as particle-deletion (due to outflow or evaporation) and list-compaction (to retain array-contiguity) are performed. This constitutes the “gather” step. In the “scatter” step to complete the communication, the global array is broadcast to all processors. Individual processors then identify their “citizen” droplets through particle-in-a-cell–based location algorithms. The drawback of this method is the cost of communicating the global buffer and the overhead resulting from the master-node-based bookkeeping. On the flip-side, gather-
scatter algorithms are robust in preventing the leakage of particles due to the occasional inability of certain processors to locate particles too close to their boundaries. In these situations, it is ensured that at least one of the abutting processors retains the particle. The point-to-point method, on the other hand, is relatively more scalable with number of processors, since the particles crossing a processor boundary are directly communicated to its new destination. All bookkeeping operations are performed locally within each processor. However, in the absence of global oversight (offered by the master processor in the collective-communication method), duplicate-particle detection is challenging in this framework.

For dense grids with more grid points than droplets, both methods offer benefits related to parallelization. In certain situations, however, both methods can fail to provide performance benefits. Consider the example shown in Fig. 5 of a charge of particles concentrated in the center prior to their convection normal to the blast wave. Obviously, the conventional domain decomposition strategy for such a highly transient problem would fail because of the location of 100% of particles in the same processor at $t = 0$, and alternative techniques such as those based on particle decomposition might be better suited for such problems. Therefore, such a method is currently being explored.

**Figure 5.** Nitromethane blast with aluminum particles (231.5 micron radius)

### 4. Applications and Postprocessing

The solver has been used for simulating turbulent combustion in realistic geometries as well as for canonical problems for a variety of applications. Applications published in recent literature include scalar mixing, extinction-reignition studies, single phase premixed and non-premixed combustion, soot modeling in premixed and non-premixed flames, spray combustion with breakup in a swirl combustor with swirler included in the computations, shock/turbulent shear layer interaction, turbulent mixing in a jet in supersonic cross-flow, and combustion modeling using real gas equation of state.

The LEM approach has been used in complex spray combustors to demonstrate its feasibility for engineering applications. For example, a through-the-vane (TTV) approach along with a breakup model for spray was used for the LDI spray combustor shown in Fig. 6(a) for modeling swirl combustion. Although the TTV approach is computationally expensive, its deployment is inevitable in situations where the boundary conditions based on frozen profiles from experimental data cannot be used just upstream of the combustor because of the penetration of the recirculating bubble upstream of the dump plane. The TTV approach is demonstrated in Fig. 6, where the CFM56 combustor's temperature field is plotted for a reacting case. As can be seen in this figure, the flame, influenced by the VBB's shape, is anchored.

Breakup models are warranted when the spray distribution is unknown close to the injector. Typically, multiple simulations are required to calibrate the inflow droplet size distributions (usually
assumed to be a log-normal). However, this approach does not guarantee that the number density distributions are matched, as shown for a simple case of a spray jet [14]. In this work, it was shown that both uniform and $1/r$ number density distributions can be calibrated to match the radial drop-size distribution. The choice of the number-density profile, however, influences the vapor distribution close to the injector, which may naturally result in inaccurate predictions for a reacting calculation. A breakup model eliminates uncertainties in droplet profiles since only the mass-flow rate of droplets and the initial droplet size (typically taken to be that of the injector diameter) are required. Because of its empirical nature, however, the breakup constants may need to be calibrated against downstream droplet distributions (or the modeler’s prior experience). This calculation can be performed in a relatively efficient manner if the simulations are rendered one-way coupled (from gas to droplets) and the velocity field is frozen (i.e., no time-integration of the gas-phase). In this manner, the cost of performing these a priori assessments is only slightly larger than that arising from postprocessing of results.

The major challenge in spray combustion is related to the cost of tracking a large number of droplets. The parcel approximation, wherein a group of particles are tracked as a single entity with identical properties, is adopted to reduce the total droplet count. Another approximation is related to the minimum size of the droplets for them to be tracked in the Lagrangian framework. The minimum size threshold increases the droplet lifetimes and therefore decreases the number of droplets tracked in the domain because of their quicker transformation to the gas-phase. In addition this also increases the Lagrangian time-step (proportional to the droplet size) thereby reducing the number of Lagrangian subiterations per Eulerian time-step. But as discussed by Srinivasan et al.[15], both assumptions need to be assessed for cost/accuracy trade-offs. In particular, choosing a high parcel density value results in the loss of accuracy in dispersion of droplets, whereas a large cutoff radius results in inaccurate predictions of the fuel vapor fields.

The challenge, as in DNS, is how to analyze and postprocess the enormous amount of data that is generated in these LES. At present, this issue remains unsolved although many strategies are being explored to process the data while the simulation is under way as well as to store data in a distributed manner. It is suggested that a major challenge for the future is not doing the actual LES of a engine-level design but understanding the results or interpreting the physics, both of which will require a new paradigm in data storage, access and processing. For example, pattern factor (the mean temperature profile at the combustor exit plane) is typically used for design analysis, and such planar information is easily obtained; but if one wishes to understand the underlying physics leading up to a specific pattern factor, then detailed analysis of the 4D (time-space) data set will be required. This is not easily feasible at this time and remains a challenge for the future.
5. Conclusions

This paper summarized some of the challenges in carrying out LES of realistic combustors and applications and also identified some of the unresolved issues in the modeling approach especially as related to the complex grid topology and two-phase applications. As summarized here, the LEM subgrid model for LES of reacting flows without any model adjustments (when combined with the LDKM closure for momentum there are no ad hoc adjustable parameters) eliminates or limits the problems with scalar closures prevalent even in conventional LES, but new issues arise with complex geometry LES. The robustness for a wide range of applications comes with an increase in cost but offers an ability to go from one simulation regime to another without changing the model—a critical strength essential to study complex flame dynamics as in combustion instability, LBO, and extinction-re-ignition processes of fundamental interest to the engine design. Currently, LES is being introduced into the design cycle at many engine developing industries (including GE, P&W, Honeywell, and Rolls Royce). However, proper establishment of LES methodology, closure, and verification of its accuracy and robustness for wide range of designs are still needed before a fully validated approach will be in place.

Acknowledgments

The results reported here are supported in part by AFOSR, ONR, DTRA, NASA, General Electric Aircraft Company, Pratt & Whitney and IBM. Simulations were conducted on our in-house PC cluster and in the Department of Defense High Performance Computing Centers at ARL (MD), ERDC (MS), NAVO (MS) and ASC (OH).

References