Table of Contents

1. Introduction .................................................................................................................. 4
   1.1 Background .............................................................................................................. 4
   1.2 Demonstration Problem ......................................................................................... 6
   1.3 Simulation Roadmap ............................................................................................... 7
   1.4 Integration .............................................................................................................. 9

2. Macroscale Team ........................................................................................................... 13
   2.1 Overview ................................................................................................................ 13
   2.2 Progress on the Demonstration Problem .............................................................. 13
   2.3 Algorithm Optimization ......................................................................................... 18
   2.4 Integration .............................................................................................................. 19
      2.4.1 Simulations of Sandia’s Multiphase Shock Tube ............................................. 19
      2.4.2 Simulations of a Randomly Distributed Fixed Particle Bed ......................... 20
      2.4.3 Simulations of Eglin microscale and mesoscale explosive experiments .......... 21
   2.5 Fifth Year Plans ..................................................................................................... 22

3. Microscale Team .......................................................................................................... 23
   3.1 Goals and Motivation ............................................................................................ 23
   3.2 Shock interaction random distribution of particles ................................................ 23
   3.3 1-D Riemann Model for Shock Particle Interaction ............................................ 26
   3.4 Shock Interaction with Moving Particles ............................................................... 26
   3.5 Shock and Contact Interaction with a 1D Array of Particles ............................... 29
   3.6 Improving PIEP Model with Regression Analysis ................................................. 33
   3.7 Fifth Year Plans ..................................................................................................... 38

4. Experiments .................................................................................................................. 39
   4.1 ASU Experiments .................................................................................................. 39
      4.1.1 Goals and Motivation ....................................................................................... 39
      4.1.2 Introduction to ASU Experiments ................................................................... 39
      4.1.3 New Equipment over 2017 ............................................................................. 39
      4.1.4 PIV Experiments ............................................................................................ 39
      4.1.5 Particle Image Velocimetry Triggered with Pressure Sensor Data .................. 42
      4.1.6 Horizontal Void Cracks ................................................................................... 42
      4.1.7 Pressure Sensor Data Examined with Differing Bed Heights ......................... 45
CENTER FOR COMPRESSIBLE MULTIPHASE TURBULENCE

7.2.4 Transferring Data ........................................................................................................ 68
7.2.5 Reinitializing Data ....................................................................................................... 68
7.2.6 Automatic Load Balancing ......................................................................................... 68
7.2.7 Experimental Results ............................................................................................... 69
7.3 DVFS on CMT-nek ......................................................................................................... 74
7.4 CMT-bone on KNL ......................................................................................................... 75
  7.4.1 KNL Architecture .................................................................................................. 75
  7.4.2 Hybrid Computing Performance (MPI + OpenMP) ............................................... 75
  7.4.3 Performance using Different Memory Modes ....................................................... 76
  7.4.4 Performance using different Clustering Modes ...................................................... 76
  7.4.5 Comparison among Different Architectures .......................................................... 77
8. Exascale Team .................................................................................................................. 79
  8.1 Overview ..................................................................................................................... 79
  8.2 Exascale Behavioral Emulation .................................................................................... 79
    8.2.1 BE – SST and its use for CMT-bone-BE simulations ........................................... 80
    8.2.2 Multi-fidelity surrogate models for performance prediction ............................... 82
    8.2.3 BE simulation of CMT-nek design space: .............................................................. 84
    8.2.4 Lab setup for power/energy/thermal modeling ...................................................... 89
    8.2.5 FPGA-acceleration methods for rapid DSE space reduction & UQ ....................... 91
9. Deep Dives ...................................................................................................................... 95
  9.1 Exascale Deep-dive ..................................................................................................... 95
  9.2 Multiphase Physics Deep-dive ..................................................................................... 97
10. Publications ..................................................................................................................... 101
11. Conferences and Presentations .................................................................................... 110
12. Workshops Held or Attended ....................................................................................... 116
13. Students and Staff Internships ..................................................................................... 117
   13.1 Internships Completed .............................................................................................. 117
   13.2 Internships Planned ................................................................................................. 119
   13.3 Internships Not Yet Planned ................................................................................... 119
   13.4 Graduated Students ................................................................................................. 119
   13.5 Placement of Staff .................................................................................................... 119
14. NNSA Laboratory Interactions ...................................................................................... 120
1. Introduction

1.1 Background

The University of Florida (UF) established a Center for Compressible Multiphase Turbulence (CCMT) on January 26, 2014 as part of the NNSA's Predictive Science Academic Alliance Program II (PSAAP-II) Single-Discipline Centers (SDC). The intellectual objectives of the Center are threefold: to radically advance the field of compressible multiphase turbulence (CMT) through rigorous first-principle multiscale modeling, to advance very large-scale predictive simulation science on present and near-future platforms, and to advance a co-design strategy that combines exascale emulation with a novel energy-constrained numerical approach. The Center is performing petascale, and working towards exascale, simulations of instabilities, turbulence and mixing in particulate-laden flows under conditions of extreme pressure and temperature to investigate fundamental problems of interest to national technological leadership. Towards this vision we are tackling the following challenges:

1) Target an important application that can only be enabled by exascale computing: We are solving a complex multiscale problem at an unprecedented level of physical detail and integration and thereby advance predictive simulation science. CMT poses a grand challenge to our understanding as it combines three complex physics: compressibility, multiphase flow and turbulence. CMT occurs often under extreme conditions of pressure and temperature, and as a result is not easily amenable to high-fidelity experiments and diagnostics. CMT presents a fascinating array of poorly-understood instability, transition, and turbulent processes manifest over a wide range of strongly interacting length and time scales. Current computational approaches involve models and closures that are developed from incomplete understanding, and as a result are largely empirical. Fully validated exascale simulation perhaps is the only path to fundamental breakthroughs that can lead us out of current empiricism.

2) Well-defined problem hierarchy leading to a demonstration problem: A multiscale approach from the microscale to the mesoscale and to the macroscale is being pursued for a systematic integrated investigation of the CMT physics. We have adopted a problem hierarchy that culminates at a signature demonstration problem of explosive dispersal of particles from a well-characterized initial condition, which fully exercises all the key complex processes of CMT. We pursue a coupling strategy where (i) fully resolved microscale simulations will lead to reduced order descriptions (interphase coupling models) to be employed at the mesoscale and (ii) partially resolved mesoscale simulations will lead to reduced order descriptions (multiphase large eddy
simulation closures) to be employed at the macroscale. This will allow computational efficiency and high degree of parallelism at all levels of the hierarchy.

3) Simulation and experiment roadmaps for rigorous validation: We focus on integrated system-scale simulations of the demonstration problem from the outset using existing integrated code capabilities. Simultaneously, we also perform petascale simulations at the micro and mesoscales. Improvements to micro-to-meso and meso-to-macro coupling models will be systematically and periodically incorporated at the appropriate higher level. A layered systems engineering approach is used to organize and integrate physical subsystems with numerical, software and service components, to achieve progressively improved operational capability for system-scale simulations. We have developed a detailed simulation and experiment roadmap which allow rigorous step-by-step validation at each step of the problem hierarchy.

4) Develop novel uncertainty quantification (UQ) approaches for CMT: Detailed measurements from carefully chosen existing and planned experiments at the Air Force Research Laboratory Munitions Directorate (AFRL-RW), Sandia Multiphase Shock Tube facility and Los Alamos Center of Mixing under Extreme Conditions (CoMuEX) are used for rigorous quantification of uncertainties from the micro/mesoscales to the macroscale. We are engaged in vigorous uncertainty reduction through better characterization and instrumentation, rigorous calibration of the models, and improved numerical resolution. Simultaneous simulations and experiments at the micro, meso and macroscales of the problem hierarchy will allow us to both propagate uncertainty to higher scales, and to reduce uncertainty through iterative improvements at the lower scales. A particularly difficult aspect of CMT is that it is characterized by extreme events that are localized in space and time. A key innovation is the development of novel techniques for accurate characterization of probability tails in the uncertainty quantification of such rare but critical events.

5) Demonstrate integrated performance on current/near-future architectures: Modern many-core architectures (such as Intel MIC), that provide high raw gigaflops, have deep memory hierarchies and low overhead threading capabilities. We exploit these capabilities to optimally utilize both computational and energy resources. In particular, we will tackle load balance and performance challenges in terms of data and work decomposition for the CMT code framework. Different parallelization schemes will be considered for effectively implementing simulations at the microscale, mesoscale, and system-scale, especially for heterogeneous resources.

6) Develop methods for predicting performance on a variety of exascale architectures: While many exascale trends seem clear, there are far too many permutations in the design space to select one a priori. We leverage the unique Novo-G facility at the NSF-supported UF Center for High-Performance Reconfigurable Computing (CHREC) to emulate and evaluate a series of candidate exascale architectures. We are developing an unprecedented capability to behaviorally prototype in software and hardware a variety of promising (as defined by leading exascale initiatives) forms of next-generation exascale (i) device and node designs at the micro-level and (ii) communication and system architectures at the macro-level. We are conducting experiments with CMT-bone
kernels, miniapps and skeleton-apps to evaluate promising architectures in terms of performance, energy, temperature, reliability, and scalability. Modeling, simulation, and estimation tools (e.g., those supported within the Sandia’s Structural Simulation Toolkit (SST)) are being leveraged with our behavioral simulations and emulations.

7) Solutions for energy efficiency and thermal management: We are developing a framework for multi-element and multi-objective optimization that will simultaneously minimize energy and maximize performance. We exploit the data and task parallelisms within CMT application and its UQ implementation to develop innovative low complexity static and dynamic algorithms for scheduling, while considering important factors such as thermal constraints and leakage currents.

1.2 Demonstration Problem

We aim at solving a problem of Compressible Multiphase Turbulence (CMT) at an unprecedented level of physical detail and thereby advance predictive simulation science. The overarching demonstration problem consists of a cylindrical core of simple explosive pellet of about 10 grams will be surrounded by a cylindrical very-thin-walled glass jacket of larger diameter. The annular region between the pellet and the jacket will be filled with mono or polydisperse metal powder of spherical shape. The shape and amount of the explosive charge and the size distribution of the metal powder and its material (aluminum, steel, tungsten, etc.) are parameters that will be varied. The charge will be hung from a test fixture so that the effect of the ground and the surrounding structures will be eliminated during the initial phase of explosion and dispersion. The orientation of the test setup will be such that the resulting explosive dispersal of particles and the gas field can be highly accurately measured. The following features makes this problem a very good choice for demonstration: (i) the explosive dispersal exercises all the major CMT physics, (ii) the extreme conditions makes this a demanding test for predictive capability, (iii) this problem requires exascale for true predictive capability, and (iv) we have already performed similar experiments.
and validation-quality measurements. The explosive dispersal of solid particles problem displayed in Figure 1.1 and described by Frost et al. (Phys. Fluids, 24(9), 2012) was chosen for the initial phase of our research activities.

1.3 Simulation Roadmap

The center is focused on integrated system-scale simulations of the demonstration problem from the outset using existing integrated-code capabilities. Figure 1.2 shows the roadmap of the proposed sequence of simulations. The following important considerations was used in constructing the roadmap: (i) Along with system-level simulations of the demonstration problem, we will perform increasingly more complex simulations at the micro and mesoscales. Based on these simulations, improvements will be made to micro-to-meso and meso-to-macro coupling

![Simulation Roadmap Diagram]

Figure 1.2. A roadmap showing the sequence of progressively more complex simulations to be performed over the five years of the center. Micro and mesoscale simulations will be performed along with macroscale simulations of the demonstration problem. Both hero runs and bundled runs will be performed. Bundled runs will be combinations of varying parameters and ensemble of realizations. This plan is based on future petascale platforms that are expected to be available through NNSA. Concurrent simulations will be accomplished using DAKOTA.
models. (ii) To take maximum advantage of validation experiments, large numbers of simulations will be required for optimal calibration. We are using surrogate models to allow us to solve the multi-level optimization problem associated with selecting the physical constants that give the best match with the numerical model. (iii) Variations of the key control parameters (particle size, particle material, shock strength, etc.) will be guided by simulations that identify which combinations of parameters will elicit different modes of instability. (iv) Statistical variability will be explored through an ensemble of realizations under nominally identical conditions. (v) Simulations are currently being carried out concurrently as *bundled runs* using the DAKOTA toolkit. (vi) We anticipate increasingly larger petascale computational platforms to be available at the NNSA labs. (vii) We have and will continue to perform selective *hero runs* at super-high resolution to help quantify discretization errors to help assess the accuracy of the estimated uncertainties. (viii) UQ is being used to guide the selections of quantities to be measured with preference to those with low uncertainty, so as to avoid empty validation based on large error bars.

The Year-1 simulations of the demonstration problem employ simplified physics model: (i) a lumped detonation model, (ii) the single-phase AUSM+ flux scheme for the Euler gas equations with idea gas equations of state, (iii) the actual particles are approximated with computational super particles, (iv) gas-particle coupling is through point-particle models of quasi-steady and unsteady forces and heat transfer, and (v) particle-particle collisions are accounted using a simplified collision model. The corresponding hero and bundled runs represent our Year-1 starting point. The above roadmap shown in Figure 1.2 lays out year-by-year progression of more detailed simulations that incorporate additional physics through new and improved models. Furthermore, each year we plan to perform larger and larger hero runs as well as large array of bundles macroscale simulations for uncertainty quantification.

The simulation roadmap is driven from the top by Uncertainty Budget (UB). A detailed phenomenon identification and ranking analysis of the demonstration problem has identified 11 key sources of errors and uncertainties which are briefly listed below:

- T1: Detonation process modeling
- T2: Multiphase turbulence modeling
- T3: Real gas thermodynamic and transport properties
- T4: Inter-particle collision modeling
- T5: Particle compaction modeling (during early stages of detonation/shock propagation)
- T6: Point particle modeling of gas-particle momentum (force) exchange
- T7: Point particle modeling of gas-particle thermal (heat-transfer) exchange
- T8: Particle deformation, sintering and break-up physics
- T9: Discretization (numerical) errors
- T10: Errors from geometric approximation (geometric differences in the details of experiments and simulations)
- T11: Experimental uncertainties and measurement errors
The key activity of UB effort will be to quantify the uncertainty in the zeroth and first order prediction metrics. The zeroth order prediction metrics of the demonstration problem are:

- The blast wave location as a function of time
- The average particle front and tail locations as a function of time
- The number of large-scale instabilities of the particulate front

The first order prediction metrics go beyond the zeroth order metrics and the details of the demonstration will be quantified with the following first order metrics:

- Time evolution of the pressure at selected points within 5% error
- Time evolution of the thermal load at selected points within 20% error
- Time evolution of average particle concentration within 15% error
- Evolution of particle front thickness due to instability and turbulent mixing within 10% error
- RMS turbulent velocity and pressure fluctuations at the particle front within 15% error,
- Time evolution of local particle size distribution within 15% error
- Multiphase turbulent spectra and correlation length scales within 20% error.

An important component of the yearly UB effort is to quantify contribution from the above 11 sources of errors and uncertainties to each of the prediction metrics. This quantification will allow us to focus on error/uncertainty reduction. Thus each year we will focus on uncertainty reduction and quantification through certain modeling and simulation activities. These are the UB drivers for the proposed roadmap and they are presented at the top row of Figure 1.2.

Figure 1.2 also presents the yearly releases of CMT-nek, the new code being co-designed through an integration of exascale higher-order algorithm with exascale emulation/simulation. Also indicated are yearly coordination with the micro, meso and macroscale experiments to be performed at Eglin Air Force Base, Arizona State University (ASU), Sandia National Laboratory (SNL) multiphase shock tube facility and Los Alamos National Laboratory (LANL) Center of Mixing Under Extreme Conditions. The macroscale simulation road map will also be supported by the yearly progression of micro and mesoscale simulations, which is also indicated in Figure 1.2.

1.4 Integration

The Center recognizes the critical importance of tight integration for the success of the center. The center will be organized in terms of tasks and cross-cutting teams, rather than in terms of faculty and their research groups. The physics-based tasks are continuous and particulates phase modeling and simulation. In addition we have exascale (EX), computer sciences (CS) and uncertainty quantification (UQ) as the cross-cutting tasks that will interface and integrate the physics-based tasks. By ensuring faculty, research scientists, and postdocs contribute to multiple physics and/or cross-cutting tasks, we will achieve tight integration. This matrix organization, depicted in Figures 1.3 and 1.4, tears down discipline and departmental boundaries and allows close interaction. In addition, significant effort has gone into integrating the various disciplines.
The intellectual integration of the different simulation and experimental talks, across the three different scales (micro, meso and macro) is shown in Figure 1.5. Uncertainty quantification, propagation and reduction along the ten sources of errors/uncertainties (T1 to T10) forms the framework that connects and drives the different simulation and experimental activities of the center. The hierarchical flow of error/uncertainty information to the macroscale is shown.
At the microscale the motion and thermal evolution of particles depends on the flow around them. In return, the particles modify the local flow by the formation of momentum and thermal wakes. Particle structures (chains and clusters) spontaneously form due to wake-wake, particle-wake and particle-particle interactions. At the mesoscale, due to inertial interaction with turbulence, particles preferentially accumulate. Also, flow instabilities can lead to large-scale structures in particle distribution. These nonuniformities have profound influence on their collective back influence on the flow. At the macroscale (or system-scale) the geometric details of the setup influence the coupling between the particles and expanding gas. Important aspects of the multiscale coupling strategy we are pursuing includes: (i) microscale-informed reduced-order descriptions (point-particle coupling models) to be employed at the mesoscale and (ii) mesoscale-informed reduced-order descriptions (multiphase LES models) to be employed at the macroscale. With this strategy, the predictive capability at the system-scale can be thoroughly validated and uncertainty rigorously quantified as illustrated in Figure 1.5.

Note that the multiscale coupling strategy and the overall uncertainty quantification plan includes both a shock-tube track and an explosive track. We have been working with the Experimental Teams at the various locations and have discussed in detail the type of characterization, inputs, and output from the experiments for a meaningful UB approach.

Finally, Figure 1.6 shows the timeline for performing the different tasks. These tasks T1-T11 were previously described.
Figure 1.6. Uncertainty Quantification Task Timeline.
2. Macroscale Team

2.1 Overview

The premier goal of the Macro/Mesoscale Team (MMST) is to run a problem of explosive dispersal of particles at an unprecedented level of details. Figure 2.1 provides an overview of the major phases of such a problem. Its second objective is to validate state-of-the-art, in-house-developed particle force models against experiments run by a number of CCMT’s collaborators. During year four, our efforts have focused on multiple targets: (i) improvement of simulations of the demonstration problem to capture more details of the phenomenon; (ii) progress on the method to maximize the late time jetting instabilities; (iii) improvement of the simulations of Eglin Air Force Base experiments; and (iv) preparation towards production-scale simulations using CMT-nek.

2.2 Progress on the Demonstration Problem

Figure 2.2 shows a snapshot of the completed Hero Run 3, along with prediction metrics 1 (blast wave location) and 2 (outer front particles location) compared with results of previous iterations of the demonstration problem. This simulation made use of a reactive burn model to generate the initial conditions for the detonation products and used an in-house-developed “single-equation...
JWL” model for the gas mixture equation of state. With an initial baseline particle volume fraction of 10%, carefully designed initial perturbations were introduced in the particle bed to gain more insight in the effect of initial conditions in problems of explosive particle dispersal. Despite this relatively low initial particle volume fraction, the particles had to be held frozen for a short period of time at the beginning of the run to prevent the particle cloud from reaching the random close packing limit. Indeed, the physics in the strong compaction/collision regime is the only element missing in this iteration of the demonstration problem. One can observe in Fig. 2.2-middle that the reactive burn initial conditions allows for an accurate prediction of the blast wave location in comparison to previous runs where a lump detonation model (along with the inappropriate ideal gas equation of state for Hero Run 1) was used. However, Fig. 2.2-right clearly indicates that freezing particles at early times modify significantly the physics of the particles dispersal to the point where predicting the particle front outer front is not feasible. Capturing the physics taking place in the strong compaction regime will be address in future runs. Also, started even before the completion of Hero Run3, the last Hero Run on the current code, Rocflu, had been started at an even greater resolution: 120 million computational cells shared on more than 16 000 processors. The next Hero Run will make use of CMT-nek.

With CMT-nek capable of handling highly compressible media and Lagrangian point-particles, a new generation of Hero runs are planned to exploit the high scalability of the code. In this direction, a preliminary simulation of explosive dispersal of particles is performed. The simulation setup is similar to that of Frost’s experiment that was used for the early Hero Runs with a ring of particles placed around a high-energy charge. For this simulation, the charge is assumed to be air at a high-pressure and density. Ideal gas law is used for the equation of state. Roughly 2500 particles are placed around the charge in a 1cm thick bed at a uniform initial volume fraction of 5%. Figure 2.3 shows the pressure field and particle curtain at t=60 µs for one these exploratory simulations. More tests and fine-tuning are ongoing in preparation for the first generation of Hero Run on CMT-nek.

Meanwhile, the optimization study of the fastest growing initial perturbation in the particle volume fraction (PVF) has continued to progress. To perform this task, the base PVF is perturbed using up to 3 sinusoidal waves with wave-numbers (k1, k2, k3). An initial hurdle was to select an objective function that would measure the jetting instabilities observed in the particle cloud. After
substantial analysis and numerical experimentation, the departure from axisymmetry in particle angular distribution was chosen. The computational domain is divided into sectors in the azimuthal coordinate and the fractional volume difference, Δf, is computed between the sector with most particles and the sector with the least particles. The fractional volume, f, is defined as the quotient between the volume of particles in a sector at time t divided by the sector volume, which in this case is a constant. Δf at time t is then divided by Δf at initial time and it is called normalized maximum PVF difference, ζ.

Figure 2.4. Overall noise in ζ was reduced from 3.2% to 0.7%. Left: ζ as a function of time. In red is plotted the original unfiltered data, in blue is plotted the filtered ζ using a 10% Fourier filter. Error bars represents one standard deviation. Right: The box plots show the overall variability in ζ(t =500µs). On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points.

Figure 2.5. Samples used to build Kriging and regression models. Left: ζ design of experiment (60 points). Right: ζ design of experiment plus permutation points (360 points).
Substantial noise is observed in $\zeta$ due to a combination of randomness in the initial position of the particles and the use of Cartesian coordinates to avoid singularities in the origin. Since a noisy function is more difficult to optimize, the noise was reduced by a Fourier filter developed in-house. Figure 2.4-left shows unfiltered and filtered $\zeta$ as a function of time. The error bars represent one standard deviation. Figure 2.4-right shows filtered and unfiltered $\zeta$ for the final time. The box edges represent the 25th and 75th percentile while the central red mark is the median. The filter reduced the noise from 3.2% to 0.7%.

If the amplitude for each of the three modes is kept the same and there is no phase shift, the order of the wave-numbers does not matter, i.e. the case with wave-numbers $(k_1, k_2, k_3)$ should have the same $\zeta$ output than its permutations $(k_1, k_3, k_2), (k_2, k_1, k_3), (k_2, k_3, k_1), (k_3, k_1, k_2), (k_3, k_2, k_1)$. Therefore, if we are dealing with costly simulations, there is no need to simulate them but just include them. We show a three-variable case for simplicity however, note that the symmetries can be imposed also in 9 variables which is the most general case.

Figure 2.5 shows the design of experiment used to train the models. Figure 2.5-left shows the original set while Figure 2.5-right shows the extended set.

<table>
<thead>
<tr>
<th>Approx.</th>
<th>RSME</th>
<th>CV-LOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta_{60,0}$</td>
<td>$1.543 \times 10^4$</td>
<td>$5.221 \times 10^3$</td>
</tr>
<tr>
<td>$\zeta_{60,360,\theta}$</td>
<td>$4.730 \times 10^3$</td>
<td>$2.763 \times 10^3$</td>
</tr>
</tbody>
</table>

Table 2.1. Left: RMSE and CV-LOO errors for Kriging surrogate using only the original 60 points ($\zeta_{60,0}$) and the one trained using 360 (60+300) points plus isotropic condition ($\zeta_{60,360,\theta}$). Right: RMSE and CV-LOO errors for linear regression surrogate using a second order degree polynomial with 10 coefficients, $\zeta_{10}$, and using 4 coefficients, $\zeta_4$.

Table 2.1-left shows the root-mean-square error (RMSE) and the Cross-Validation leave-one-out (CV-LOO) errors for the two studied training cases using Kriging surrogate. $\zeta_{60,0}$ represents the approximation using the original 60 points and $\zeta_{60,360,\theta}$ using the permutation points plus isotropic Kriging (i.e. using the same hyperparameter for all the variables). The RMSE improved by 69% while the CV-LOO improved by 47%. Table 1-right shows the RMSE error and the CV-LOO errors for the two studied training cases using linear regression. $\zeta_{10}$ represents the approximation using a second order degree polynomial with 10 coefficients. $\zeta_4$ represents the approximation using also a second order degree polynomial but imposing the symmetries of the problem modifying the basis functions as follows $a_0 + a_1 (k_1 + k_2 + k_3) + a_2 (k_1^2 + k_2^2 + k_3^2) + a_3 (k_1 k_2 + k_2 k_3 + k_3 k_1)$ where $a_0, a_1, a_2$ and $a_3$ are the polynomial coefficients. While the CV-LOO remains similar for the approximations, the RMSE improved by 37%.
One last important achievement for the MMST with regards to the demonstration problem was to initiate simulations of the Eglin blastpad experiments illustrated in Fig. 2.6. The first simulations were set up in a one-dimensional domain extending 6 m. radially from the center of the Comp B charge. The JWL equations were used as the equation of state for the charge, which was assumed to instantaneously convert into its gaseous detonation products as the detonation starts. The data obtained from these runs allowed for the tracking of the blast wave, detonation product contact interface and the secondary shock as a function of time. These results were shared with the experimental group at Eglin for comparison with new and legacy bare charge experimental data.

Next, 2D simulations were performed. The computational domain is shown in Fig 2.7-left. It consists of a smaller square section which represents the cavity in the ground where the charge is placed in the experiments (see Fig. 2.6). The top part of the computational domain represents the outside ambient space. Overall, the grid counts about one million computational cells. The dimensions of the ambient were chosen to allow for the two innermost E-line experimental probes to lie inside of the domain while allowing for a relatively small-scale run and ensure that the simulation is behaving as expected before we transition to a full-scale, 3D version. A reactive burn

Figure 2.6. Charge suspended before a blastpad experiment at Eglin Air Force Base.

Figure 2.7. Left: 2D computational domain for the blastpad experiment. Orange dots are pressure probes that match probe locations in the experiment. Black dots are additional pressure probes near the charge for detailed analysis. Right: Pressure recorded in the 2D simulation by two of the pressure probes that match experimental probes’ location.
profile for the charge’s initial state is in preparation for the next series of tests. This is being done using an in-house code which simulates the burning of the Comp-B charge in any given two-dimensional domain by solving the reactive Euler equations.

2.3 Algorithm Optimization
The MMST-UB Team joint efforts for the development of Kriging surrogate models for gas mixture equation of state has been finalized. These models replace the iterative solver for the nonlinear system of equations which arise from the evaluation of the equation of state for a mixture of the gaseous detonation products of PETN and air. They were created using different numbers of sampling points ranging from 50 to 800 to quantify the effects of refinement on the model behavior. Selected results of the initial testing of these models for their computational efficiency and accuracy with respect to the iterative solver are shown in Fig. 8. Figure 2.8-left shows that the use of the Kriging models leads to a decrease in the amount of computing time required for a hydrocode to perform a simulation involving a PETN product and air mixture. Further analysis also indicates that models created using more than 200 sampling points perform well in terms of tracking the blast wave of the detonation (not presented here). This leads to the conclusion that there appears to be a good middle ground between computational speed and model accuracy, illustrated in Fig. 2.8-right, with models generated using 200 sampling points. These results pave the way for further development for mixtures of more than two gases or equations of state. This is an area where the potential for improvement in computational time is high.

Figure 2.8. Left: Simulations run with each of the mixture surrogates (50-point, 100-point, 200-point, 400-point, and 800 points models) along with associated speed-up ratios with respect to the iterative solver for gas mixture. The error bars represent one standard deviation of the averaged data sets. Right: Pressure profiles for the PETN detonation simulations using the iterative EoS solver and the 200-point surrogate model at given times.
2.4 Integration

2.4.1 Simulations of Sandia’s Multiphase Shock Tube

A new set of 1D, 2D and 3D simulations (Fig 2.9-top) of Sandia’s Multiphase Shocktube experiments were performed with the initial volume fraction profile modeled using a Radial Basis Neural Network (RBNN) as seen on Fig. 2.9-bottom-left. This profile was modeled closely to match the shape of the curtain obtained from experiments. A combined total of 79 (45 1D, 25 2D & 9 3D) simulations are performed for an initial curtain thickness in the range of 2mm to 3mm at volume fractions ranging from 19%-23%. While a few simulations are left to run, analysis of the completed runs seem to indicate that 2D and 3D simulations do not agree on the trajectory of the particles curtain downstream front. In contrast, predictions of the upstream front are consistent. A detailed uncertainty budget is being conducted by the UB team and will help determine the origin of this discrepancy.

Figure 2.9. Top: Perspective view (top), side view (middle) and top view (bottom) of a 3D simulation of Sandia’s Multiphase Shocktube experiments at $t \approx 350\mu s$. Computational particles are colored by their velocity magnitude, the back wall is colored by the magnitude of the density gradient (gray scale) and the bottom wall is colored by pressure. Bottom-left: initial particle volume fraction profile modeled using a Radial Basis Neural Network. Bottom-right: particles curtain upstream front and downstream front trajectories from the experiment compared with trajectories extracted from 2D and 3D simulations.
2.4.2 Simulations of a Randomly Distributed Fixed Particle Bed

Next, the ongoing effort to develop better force models at higher Mach numbers information from microscale simulations is being used to drive mesoscale simulations. For this, simulations of a Mach 3.0 shock going over a fixed bed of particles are performed at volume fractions of 10%, 15%, 20% and 25%. In all the simulations, particles of size 100 µm in diameter \(d_p\) are populated in a 2mm \(L_y \times L_z\) size curtain, where \(L_y\) and \(L_z\) are the dimensions of the bounding box in cross-stream directions. The total number of Lagrangian point-particles in these simulations are chosen such that the super-particle loading remains constant. This corresponds to 2, 3, 4 and 5 particles per cell for 10, 15, 20 and 25% volume fraction cases respectively. The mesoscale simulations replicate the particle-resolved microscale simulations. Mean pressure profiles are shown at two different volume fractions in Fig. 2.10. The \(x\)-coordinate is normalized by the particle diameter and the \(y\)-coordinate is normalized by the post-shock pressure. The mesoscale simulations capture the peak quantities to a good extent as shown in Fig. 2.10. However, the shock front position is slightly behind that of microscale simulations. Clearly, the drag models are predicting a higher drag in the curtain resulting in the discrepancy of the location of transmitted and reflected shock fronts. This is amplified at higher volume fractions. Since the drag models are developed for a single particle, effects of the neighboring particles and particle wake is not captured in these models. Also, at \(M=3.0\), a bow-shock forms in the microscale simulations whose effects are not captured in the current models.

![Figure 2.10. Planar averaged pressure profiles normalized by post-shock pressure at volume fraction \(\phi_p = 10\%\) (left) and \(20\%\) (right) are plotted against the streamwise spatial co-ordinate \((x/d_p)\) at \(t/\tau = 0\) (red), \(t/\tau = 4\) (blue), \(t/\tau = 8\) (blue), \(t/\tau = 12\) (black) where \(\tau = d_p/u_{\text{shock}}\).](image-url)
2.4.3 Simulations of Eglin microscale and mesoscale explosive experiments

Simulations of the Eglin microscale and mesoscale explosive experiments have been continuously refined throughout year four with the help of the experimental team to fully capture the subtleties of the initial conditions of these experiments. Indeed, the detonator access hole at the back of the casing has now been included in the computational domain (see red area in Fig. 2.11). Also, a reactive burn model for the explosive detonation products has been computed and implemented. Finally, explosive initial conditions as well as geometric features have been and continue to be fine-tuned to emulate the effects of the casing deformation observed in some experiments. The fast pace of the recent improvements in predictions by these simulations has been possible since the implementation of the Navier-Stokes Characteristic Boundary Condition to handle the outflow of explosive products while using the JWL equation of state. This feature has allowed for smaller computational domain; thus, it has helped in reducing the computational time required to carry out the JWL simulations of this problem. Figure 2.12 shows the gas density gradient contours and the particle position at \( t = 20\mu s \) for a recent simulation. In this run, the combined effect of a slightly modified explosive, the reactive burn model initial conditions, a parabolic initial explosive front, and a conical barrel shape has significantly improved the predicted trajectories of particles in comparison with the past microscale experiments as illustrated in Fig. 2.13. As this work on the microscale experiments presses on, simulations of the mesoscale experiments are being carried out with the processing of the experimental data for the particle cloud.

![Figure 2.11. Eglin Air Force base microscale experiment configuration.](image)

![Figure 2.12. Axisymmetric simulation of Eglin AFB’s microscale experiment. The gas density gradient contour and particle position at \( t = 20\mu s \). The particle originates from the barrel exit and the simulation time begins the moment when the detonation wave has propagated through all explosives. Note that the particle is oversized for ease of visualization.](image)
2.5 Fifth Year Plans

For year Five, the MMST is fully transitioning to CMT-nek. The center’s developed code has now sufficient capabilities to perform meaningful simulations of most of the experiments that the Macro/Mesoscale Team were running with Rocflu. The one major objective for the team is to run full scale simulations of the demonstration problem.

In parallel, the team has started drafting numerous archival papers with the results obtained throughout year four. These papers will be submitted for publication in year five.
3. Microscale Team

3.1 Goals and Motivation

Goals of Microscale team include performing fully resolved simulations of shock and contact interacting with bed of particles. Purpose of these simulations is to understand the underlying complex physical mechanisms occurring during blast conditions and converting this understanding into models, which can be used to predict the particle motion and back effect of particles on the flow.

3.2 Shock interaction random distribution of particles

Fully resolved three-dimensional inviscid simulations of shock interacting with particles were carried out. The particle volume fraction varied from 10% to 25% and number of particles in the domain varied from 200 to 500. During the early times that we are interested in, particle motion and viscous effects are negligible. Hence, we solve Euler equations on a body fitted grid with stationary particles. Simulation setup is shown in Figure 3.1.

To understand the effect of the incident shock on the particles, we plot the non-dimensional streamwise drag force, $C_D$, along with the peak streamwise drag force, $C_{D, peak}$. The non-dimensional drag force is normalized by the peak streamwise drag induced by the incident shock, $C_{D, peak}$. Results for $M_s = 3.0$ and $\phi_1 = 2.5, 10$ and 20% are plotted in Fig. 3.2. It is interesting to observe that some of the particles have $C_D/C_{D, peak}$ greater than 1.0. This indicates that drag force at later times can be more than that induced by the incident/transmitted shock, highlighting the effect of the secondary wave interactions.

We plot the least squares fit (red curve) for the $C_{D, peak}$ data. It is clear from this plot that the mean peak drag force decreases as the particle bed depth increases. This indicates that the transmitted shock weakens as it travels through the particle bed. It is clear from the drag plot that there is
significant fluctuation in the drag force seen by each particle. Each particle diffracts the imposed flow field that it sees, and random distribution of particles results in unique flow field that each particle sees. This causes significant variation in the drag force. Nonlinear effects and fluid
mediated particle-particle interaction also results in transverse forces on the particles. The transverse forces are about $20 - 50\%$ in magnitude compared to the streamwise drag force.
Current point particle models do not consider the effect of transverse forces as well as the particle to particle fluctuations in the drag force. Based on the knowledge from the fully resolved simulations, it is clear that there is a need to improve the current point particle models.

3.3 1-D Riemann Model for Shock Particle Interaction

Shock interaction with bed of particles results in complicated wave dynamics, which can include reflected shock, transmitted shock, expansion fan, and contact interface. There is also possibility of resonance, when to wave families coincide with each other. To better understand this weave dynamics, we pose the problem of shock interaction with bed particles in 1-D domain. In this 1-D domain particle bed is treated as region with area change. This is a Riemann problem with discontinuous area change. We obtain the solution to this problem for different area changes (10%-25%) and $M_s = 3.0$. Wave structure from the Riemann solution is shown in Figure 3.3. Flow field resulting from Riemann solution is plotted in Fig. 3.4.

3.4 Shock Interaction with Moving Particles

Fully resolved three-dimensional simulations of shock interacting with particles were carried out. Multiple simulations were carried about varying the particle volume fraction, shock Mach number, and particle to gas density. Time scale for the motion of particle is directly proportional to particle to gas density ratio. Results from one such simulation will be discussed in this report. Incident shock with Mach number, $M_s = 3.0$, was initialized upstream of a bed of particles having particle volume fraction of 10%. Particle to gas density ratio was 530. There were 100 spherical particles in the computational domain.

In Figure 3.5 non-dimensional streamwise drag force, $C_D$ experienced by every particle is represented along with plot of velocity magnitude as function of non-dimensional time, $\tau_s$. Where, $C_D =$

![Figure 3.5. Plot of (a) non-dimensional drag force in streamwise direction, $C_D$, against non-dimensional time, $\tau_s$, and (b) velocity magnitude against non-dimensional time $\tau_s$.](image-url)
$F/(0.5\rho_2 u_2^2 A_p)$, and $\tau_s = (U_s t)/D_p$. It is clear from these two plots that there is lot of variability, and each particle experiences a slightly different flow field, which is further amplified as the particles move.

As seen from Figure 3.5, when results from all the 100 particles are presented on a single plot, it gets difficult to isolate certain particles and observe the finer details. To overcome this, particles are sorted into different ‘bins’ based on their streamwise location at time zero. Particles are tagged based on their ‘bin number’, and tracked over time. In this analysis we have considered 4 bins of equal width. At time zero particles are randomly distributed between $x = 0$ and $x = 8D_p$ in the streamwise direction. For 4 bins, each bin is $2D_p$ in width. Bin1 goes from $x = 0$ to $x = 2D_p$, Bin2 extends from $x = 2D_p$ to $x = 4D_p$ and so on and so forth.

Plot of CD against $\tau_s$ is presented in Figure 3.6 (note that the time scale is zoomed in to focus on early time behavior), and (on average) peak drag (Peak drag due to the shock) goes down for
subsequent bins. This highlights the loss in strength of the incident shock wave as it travels through the particle bed. Another interesting feature is that the long term drag (after 30 shock time scales) is significantly higher for bin3 and bin4 than bin1 and bin2. This indicates that bin3 and bin4 are going to have higher velocities around that time and much more rapid transport. These observations are consistent with our previous study of shock interacting with stationary particles.

Streamwise velocity as function of non-dimensional shock time scale is plotted in Figure 3.7. There is a clear trend that average streamwise velocity increases as bin number increases. In other words, streamwise velocity increases as we travel deeper into the particle bed and close to the rear edge. This observation is consistent with the force observation that particles in bin3 and bin4 move faster and farther than particles in bin1 and bin2.

We have computed the projected area of the particles as function of streamwise location, over the simulation time. Projected area can be visualized as the amount of blockage particles provided for the flow in streamwise direction. If the projected area is high at any location, then that means there are high number of particles present in that location and low projected area means less number of
particles. Results at four various times are presented in Figure 3.8. Projected area is normalized by the cross-sectional area of 1 particle. Time in Figure 3.8 is in micro seconds. It is evident from Figure 3.8 (b)-(d) that there is a distinct signature (spikes indicating large concentration of particles) of particles forming clusters which translate in the streamwise direction. One crucial factor to be considered here is that of particles leaving the domain (once they leave the domain their area is not considered). Since number of particles is not conserved, magnitude of the spikes in the area plots keeps decreasing with time.

3.5 Shock and Contact Interaction with a 1D Array of Particles

Shock interaction with particles has been studied extensively; however, very little is known about the effects a contact discontinuity has on particles. To characterize the interaction of a shock and contact discontinuity with a particle bed, several fully-resolved, inviscid simulations were performed. Analysis of the simulations was primarily focused on particle force histories, behavior of the fluid in the particle bed, and the phenomena present with varying particle volume fraction and Mach number.

The simulation domain consists of ten stationary particles arranged in a simple cubic array and a shock and contact initialized near the leading edge of the first particle, an arrangement analogous
to an array of particles situated near the diaphragm of a shock tube. A diagram of the initial conditions is shown in Figure 3.9.

The simulations are divided into the following sets based on initial conditions: close, intermediate, and shock only. “Close” denotes that the initial spacing between the shock and contact is small (about an eighth of a particle diameter) so that the particles experience a combined presence of the shock and contact. “Intermediate” denotes that there is an intermediate separating distance (about 1.6 particle diameters) between the initial position of the shock and contact so that individual effects of each discontinuity on force history are more prominent. “Shock only” denotes the lack of a contact and serves as a means of comparison to determine the impact of the contact on particle force history. The volume fraction as well as the shock and contact Mach numbers were varied to determine how they affect shock-contact-particle interaction. The matrix of simulations is shown in Table 3.1.

Table 3.1: Matrix of simulations for shock followed by contact over a simple cubic array. Note that there is no contact for the Shock Only runs so the $M_{ct}$ does not apply to that section.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$M_{ct} = 0.31, M_s = 1.22$</th>
<th>$M_{ct} = 0.90, M_s = 1.90$</th>
<th>$M_{ct} = 1.26, M_s = 2.69$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>RUN1</td>
<td>RUN2</td>
<td>RUN3</td>
</tr>
<tr>
<td>10%</td>
<td>RUN4</td>
<td>RUN5</td>
<td>RUN6</td>
</tr>
<tr>
<td>20%</td>
<td>RUN7</td>
<td>RUN8</td>
<td>RUN9</td>
</tr>
<tr>
<td>40%</td>
<td>RUN10</td>
<td>RUN11</td>
<td>RUN12</td>
</tr>
<tr>
<td>5%</td>
<td>RUN13</td>
<td>RUN14</td>
<td>RUN15</td>
</tr>
<tr>
<td>10%</td>
<td>RUN16</td>
<td>RUN17</td>
<td>RUN18</td>
</tr>
<tr>
<td>20%</td>
<td>RUN19</td>
<td>RUN20</td>
<td>RUN21</td>
</tr>
<tr>
<td>40%</td>
<td>RUN22</td>
<td>RUN23</td>
<td>RUN24</td>
</tr>
<tr>
<td>5%</td>
<td>RUN25</td>
<td>RUN26</td>
<td>RUN27</td>
</tr>
<tr>
<td>10%</td>
<td>RUN28</td>
<td>RUN29</td>
<td>RUN30</td>
</tr>
<tr>
<td>20%</td>
<td>RUN31</td>
<td>RUN32</td>
<td>RUN33</td>
</tr>
<tr>
<td>40%</td>
<td>RUN34</td>
<td>RUN35</td>
<td>RUN36</td>
</tr>
</tbody>
</table>
All simulations were performed on Quartz and have completed. Figure 3.10 displays the force history on ten particles for RUN6, RUN18, and RUN30, a comparable simulation from each set. There are distinct maxima due to the contact on each particle in (b) and (c) that are not present in (a) that decrease with each successive particle. It is easier to distinguish between the shock and contact peak forces in the intermediate case due to the initial separating distance of the discontinuities. Figure 3.11 shows a comparison of force histories with varying volume fraction from RUN3, RUN6, RUN9, and RUN12. From 5% to 20% volume fraction, the leading particle’s sustained force decreases more rapidly as volume fraction increases and peak force remains steady; however, at 40% volume fraction the first particle experiences a significant increase in peak force as well as a higher sustained force when a reflected wave develops and travels upstream.

The shock shows a tendency toward a planar interface when deflected but the contact becomes mixed very rapidly and no planar interface is distinguishable; thus, many parameters are being explored to track the location of the contact and investigate the phenomena within the flow field. Figure 3.12 reveals four such parameters for the same time step of RUN6. The leading interface in each is the shock; the contact is much more difficult to observe. The first two parameters are pressure and Mach number. There are higher pressure regions behind the shock up to about the 5th particle that indicate a disturbance, but there is no clear indication of where the contact has reached.
The Mach contour indicates there are expectedly low velocity regions between the particles and high velocity regions between the particle and the wall. The third contour displays the density field and the fourth displays the density gradient. The density gradient most adequately reveals the presence of discontinuities. The shock as well as reflected waves and the formation of bow shocks are plainly visible.

Figure 3.13 shows the density gradient magnitude of RUN4, RUN5, and RUN6. Figure 3.13a shows the case when post-shock flow is subsonic; the contact interface has become highly distorted/stretched with the formation of a wake that lengthens as it traverses the particle bed. The many reflected waves formed as the primary shock passes through the particle bed are also clear. Figure 3.13b shows the case when post-shock flow is nearly supersonic. Regions of high pressure form very rapidly at the leading edge of the particles and a pressure wave is formed that travels upstream. Shocks form extending from particle surface to the domain boundary and the formation of reflected waves from the primary shock is clearly visible on the ninth particle. The contact becomes rapidly mixed resulting in a lack of a distinguishable interface; however, the region up to the sixth particle contains a mixture of high and low gradients and is indicative of the presence of the mixed contact. The final case (Figure 3.13c) corresponds to supersonic post-shock flow. The
formation of bow shocks at the leading edge as well as oblique shocks as the fluid passes each particle are clearly visible. The contact is again rapidly mixed so that the interface is indistinguishable but it also contains a region of high and low gradients to indicate the presence of the contact.

3.6 Improving PIEP Model with Regression Analysis

The development and improvement of a computationally efficient point-particle model is one of the objectives of the Microscale team. In previous works, the introduction of the Pairwise-Interaction Extended Point-Particle (PIEP) model has given us the ability to account for fluid mediated particle-particle interactions that lie in the unresolved length-scale of Euler-Lagrange simulations. While the PIEP model is a powerful tool, it makes several assumptions in its use of the generalized Faxén’s theorem, and these assumptions leads to increased error at higher particle volume fractions. To extend the applicable range of the PIEP model to high volume fractions, a data-driven approach is utilized.
In previous work, the PIEP model for the hydrodynamic force and torque on a particle due to its neighboring particles has been developed. This model is expressed its force as the summation of the quasi-steady, undisturbed, added-mass, Basset-history, lift, and Faxén form correction forces induced by neighboring particles. This PIEP model was derived using a theoretical understanding of the physics involved and is therefore referred to as the physical PIEP model or “PIEP-Phys.” The force in this existing model is given as the sum of the quasi-steady, undisturbed, added mass, basset history, lift, and Faxén form correction forces as

\[ \mathbf{F}_{\text{Hyd,PIEP}} = \mathbf{F}_{qs} + \mathbf{F}_{un} + \mathbf{F}_{am} + \mathbf{F}_{vu} + \mathbf{F}_L + \mathbf{F}_c \]

Similarly, torque can be written as the sum of the quasi-steady, undisturbed, Faxén form correction, and statistical volume fraction torques as

\[ \mathbf{T}_{\text{Hyd,PIEP}} = \mathbf{T}_{qs} + \mathbf{T}_{un} + \mathbf{T}_c + \mathbf{T}_\varphi \]

In this work, we introduce two more forms of the PIEP model. The first is derived completely from direct numerical simulation (DNS) data and will be referred to as the data-driven PIEP model. The next model is one that combines the physical PIEP model along with the data-driven methods to
Center for Compressible Multiphase Turbulence

construct a hybrid data-driven, physics-informed model. This model is referred to as the hybrid model or “PIEP-Hyb”. To construct the hybrid model, a statistical volume fraction force, \( F_\varphi \), is added to the PIEP model. This new summation is given as

\[
F_{\text{Hyd,PIEP}} = F_{qs} + F_{un} + F_{am} + F_{vu} + F_L + F_c + F_\varphi
\]

Similarly, torque can be written as the sum of the quasi-steady, undisturbed, Faxén form correction, and statistical volume fraction torques as

\[
T_{\text{Hyd,PIEP-Hyb}} = T_{qs} + T_{un} + T_c + T_\varphi
\]

To develop the data-driven components of the models, DNS data was acquired for a random array of monodispersed particles. The DNS data was that produced by Akiki et al (2016). In that study, a random array of equally seized particles was distributed in a cubic domain with uniform probability. Periodic boundary conditions were implemented in the flow direction (x-direction) and a direction normal to the flow (y-direction). A no-stress boundary condition was implemented at the boundaries in the z-direction. The fluid was fully resolved around each particle by solving the incompressible Navier-Stokes equations with no-slip and no-penetration boundary conditions on each particle’s surface. These no-slip and no-penetration conditions were satisfied using the immersed boundary method. The particles’ time-averaged drag, lift, and torque values were obtained. To avoid the effects of the boundary conditions in the z-direction, only particles from the middle 64% percent of the domain (in the z-directions) were considered for force modeling. This DNS was performed at various particle volume fractions and Reynolds numbers. Finally, data was separated into two categories, the regression and post-regression data. The regression data was used to train the statistical model, and the post regression data was used to determine the predictive capabilities of the model. The DNS seed that was selected to be the post-regression data was selected at random.

The models were then trained using MATLAB’s nonlinear regression function. This function takes the following as inputs: predictor values, response values, functional form, and initial parameter guesses. It then uses the Levenberg-Marquardt least squares algorithm to interactively pick the parameters of the model. The minimum mean squared error that a regression can obtain indicates how well the functional form can related the predictor values to the response values. For this reason, it is important to pick a functional form that can describe the relationship between the predictor variables and response.

For this model, the predictor variables were chosen to be the relative locations of neighboring particles to a particle of interest, and the response variable was the error in current PIEP model’s predictions. The model’s functional form was then selected to be a linear combination of Neumann and Bessel functions multiplied by the associated Legendre function. This functional for was motivated by the axisymmetric solutions to the Helmholtz equation in spherical coordinates. This
function was selected over others due to its performance with regards to the post-regression data. The metric used to select this function was the coefficient of determination, denoted as $R^2$, of the model’s predictions vs the DNS data. $R^2$ can be expressed mathematically as

$$R^2 = 1 - \frac{\sum_{n=1}^{N_y} \left[ F_{y,DNS}(n) - F_{y,PIEP}(n) \right]^2}{\sum_{n=1}^{N_y} \left[ F_{y,DNS}(n) - \langle F_{y,DNS} \rangle \right]^2}.$$

Though the above expression is for the force in the y-direction (lift), it can also be extended to the force in the stream-wise/x-direction (drag) or torque.

Once the new model was trained, it was tested against the post-regression data to determine its ability to produce accurate predictions. Table 3.2 reports the $R^2$ values for the force stream-wise direction (drag), force in the y-direction (lift), and torque in the z-direction. It also includes the previous PIEP model’s (without the data-driven term) $R^2$ values for the sake of comparison. It should be noted that these $R^2$ values will differ slightly depending on the seed selected to use as the post-regression data. Therefore, the post-regression data was selected at random.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>Re</th>
<th>Drag</th>
<th>Lift</th>
<th>Torque</th>
<th>Drag</th>
<th>Lift</th>
<th>Torque</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>40</td>
<td>0.70</td>
<td>0.68</td>
<td>0.75</td>
<td>0.72</td>
<td>0.75</td>
<td>0.79</td>
</tr>
<tr>
<td>10%</td>
<td>70</td>
<td>0.65</td>
<td>0.68</td>
<td>0.65</td>
<td>0.67</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>10%</td>
<td>173</td>
<td>0.35</td>
<td>0.59</td>
<td>0.45</td>
<td>0.44</td>
<td>0.60</td>
<td>0.58</td>
</tr>
<tr>
<td>20%</td>
<td>16</td>
<td>0.38</td>
<td>0.34</td>
<td>0.48</td>
<td>0.66</td>
<td>0.74</td>
<td>0.71</td>
</tr>
<tr>
<td>20%</td>
<td>89</td>
<td>0.50</td>
<td>0.48</td>
<td>0.72</td>
<td>0.62</td>
<td>0.63</td>
<td>0.77</td>
</tr>
<tr>
<td>45%</td>
<td>21</td>
<td>0.01</td>
<td>0.09</td>
<td>0.47</td>
<td>0.55</td>
<td>0.59</td>
<td>0.76</td>
</tr>
<tr>
<td>45%</td>
<td>115</td>
<td>0.21</td>
<td>0.19</td>
<td>0.51</td>
<td>0.63</td>
<td>0.57</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Table 3.2: Resulting $R^2$ Values for the PIEP Model with and without a Data-Driven Component.

As expected, the addition of a data-driven component drastically improved the PIEP model’s coefficient of determination at higher volume fractions. At the volume fraction of 10%, the physical PIEP model’s predictions were only slightly enhanced by the addition of the data-driven term due to (1) the already high accuracy of the physical PIEP model at low volume fractions and (2) the lack of training data due to the low volume fraction. The lack of training data at low volume fractions could also explain the purely statistical model’s low $R^2$ at this volume fraction. The use of the hybrid formulation allows the PIEP model to utilize the strengths of the physical approach at low volume fractions and the statistical approach at high volume fractions.
Now, we will further investigate the results of the three models. In way of example, we will only consider the case of \( \varphi = 45\% \) and \( Re = 21 \) where the data-driven and hybrid models perform considerably better than the physical model. Figure 3.14 shows contour plots of drag, lift, and torque maps plotted on the streamwise (x-axis) and transverse (y-axis) plane for the three approaches: physical, data-driven, and hybrid. These maps must be interpreted as follows. Given the reference particle (marked by the black circle) located at the center of the frame, the contour plots in the top row of frames show the contribution to drag force from the reference particle on a neighbor whose center is located at a point \((x, y)\) relative to the reference particle. Since the neighbor's center cannot get closer than one particle diameter, we see a white annular band of excluded region around the reference particle. Similar interpretations apply to lift force and torque contours as well.

We first note that the left column of contours corresponding to the physics-driven approach are very similar to those for the lower volume fraction. This is due to the fact the physics-driven maps are based on DNS of flow over an isolated particle with correction from flow over a pair of particles. Thus, these maps depend only on Reynolds number and do not vary with volume fraction. This is a fundamental limitation of the physics-driven approach. The average flow through the particle array changes substantially with increasing particle volume fraction and this effect is not properly reflected in the wake of an isolated or a pair of particles. For this reason, the hybrid model becomes more similar to the data-driven model as the volume fraction increases since the physical model is not accurate at high volume fractions. At lower volume fractions, the hybrid model will be more like the physical model.
The effect of volume fraction is well captured in the data-driven approach, since the actual DNS data is being used in obtaining the PIEP-data force and torque maps. Comparing the physical and data-driven PIEP models, two interesting observations can be made. First, with increasing volume fraction the contours have shrunk and gotten close to the reference particle, which is consistent with the fact that at $\varphi = 45\%$ the average distance between the reference particle and its neighbors has considerably reduced. Second, the maps show a much higher degree of left-right symmetry at the higher volume fraction. In other words, it appears that the long wake on the leeward side of the particle that breaks the left-right symmetry at low volume fraction is not an appropriate model at higher volume fraction. It is interesting to note that at higher volume fraction the average flow regains a higher level of symmetry, and thereby becomes more organized.

To summarize: using a combination of fluid dynamics and statistics, the hydrodynamic interactions between neighboring particles were accurately approximated. The PIEP model is currently one of the best models for approximating interactions of particles in a particle cloud, but the PIEP model also struggles to approximate drag or lift at higher volume fractions. In order fix these issues, nonlinear regression was utilized. Using a functional form inspired by the Helmholtz equations, the hydrodynamic interactions between neighboring particles were described statistically. In the cases where the physical PIEP model was not able to achieve a high coefficient of correlation, the statistical model was added to the physical PIEP model to create the hybrid PIEP model. This improvement extends to high Reynolds numbers as well. The results of the regression can now be examined to better understand the phenomena behind particle interactions. Work has already begun to implement these statistical models. By interpolating the model's resulting maps, accurate approximations of particle interactions can be made for Reynolds numbers and volume fractions between those used to train the models. Since this paper yielded positive results, we are also looking to implement more cutting edge algorithms, such as various types of neural networks.

3.7 Fifth Year Plans

Simulation matrices not completed in year 4 will be completed in year 5. Numerous archival quality papers are currently being written, and year 5 will see all year 4 work submitted for publication in archival journals. We will also continue our effort at developing point-particle force models that can be used at the macroscale. In addition, year 5 will see incorporation of PIEP models into CMT-nek.
4. Experiments

4.1 ASU Experiments

4.1.1 Goals and Motivation

Turbulent multiphase flows are highly complex. The simulation team at the University of Florida is working to simulate an exploding cylindrical charge in its entirety. This multifaced problem has been broken down into several smaller segments, including the decompression of a densely-packed particle bed.

The experimental team at Arizona State University is working to gather experimental data on the decompression of a densely-packed bed to help validate the early stage codes being developed at the University of Florida.

A particle bed composed of small (< 1mm diameter) glass beads is placed in the bottom of a vertical shock tube. A diaphragm is placed above the particle bed and everything above the diaphragm is vacuumed down to a low pressure \( p_1 < 20 \text{kPa} \). Everything below the diaphragm remains at atmospheric pressure \( p_4 = 101.3 \text{kPa} \). Once the diaphragm is ruptured, the pressure differential causes material (air and particles) to rapidly accelerate. The time frame of interest during each realization is on the order of milliseconds.

4.1.2 Introduction to ASU Experiments

The Laboratory for Energetic Flow and Turbulence at Arizona State University is using a vertical shock tube to examine the expansion fan behind a shock wave. The expansion fan travels into the high pressure region, which contains a particle bed of small glass beads. The shock wave will travel upward into the low pressure region. High-speed video, pressure sensor data, and PIV data is recorded.

4.1.3 New Equipment over 2017

The new pressure sensors and data acquisition device arrived at ASU’s experimental facility. With this new setup, the team is able to obtain pressure data with an appropriate temporal resolution required to see the dynamic pressure changes, including shocks and reflected shocks. Additionally, with the new computer, we are able to synchronize the pressure sensor data and the video data. As of April 2017, all the timing cards, data acquisition cards, and data acquisition devices are connected and working properly.

4.1.4 PIV Experiments

In order to explore the gas dynamics in the shock tube, we setup particle image velocimetry (PIV) experiments. Our PIV experiments use an illuminated planar cross-section (approximately 2mm thick) of the flow with a Litron Nano L 200-15 pulsed Nd:YAG laser. The laser is synchronized with a high-speed camera so the flow is sufficiently illuminated when images are collected. As the laser illuminates both glass particles (from the bed) and tracer particles (passively following
the gas flow), we use 10 micron ($D_{50} = 10 \mu m$) silver-coated hollow glass spheres to serve as passive gas tracer particles. These tracer particles are sufficiently smaller than the O(100 micron) glass particles, so we are able to differentiate between the gas tracer particles and glass particles by using a size threshold in the PIV analysis.

A possible change to future PIV experiments could include using fluorescent particles with wavelength-filtered cameras to better distinguish between the particles. An additional method we are considering would be using smoke to trace the gas flow, similar to preliminary experiments that were performed in Fall 2016. By using smoke, we could perform techniques based on laser induced fluorometry (LIF) to evaluate the evolution of the gas concentration as it compresses and decompresses along the shock tube.

By using these optical techniques, we have several main objectives to accomplish. Using a vertical light sheet that passes through the center of the tube allows for 2-dimensional 2-component velocity measurements of both the gas and dilute particle bed. It is critical not to perform these measurements in a dense particle bed, as neither the laser light sheet nor camera will be able to penetrate through the measurement region of interest. Thus, these measurements are performed above the bed. A schematic of this setup is shown in Figure 4.1. With resulting PIV analysis of glass particle and gas tracer particle velocity fields, we will be able to quantify the transition to turbulence and relationship between gas and particle motions. Similarly, in this setup, if we use

![Figure 4.1: Laser and optics produce vertical light sheet, illuminating both glass particles and gas tracer particles passing through the center of the tube. The camera records images of particles passing through field of view for PIV analysis of vertical and radial gas and particle motions.](image1)

![Figure 4.2: Laser and optics produce horizontal light sheet, illuminating glass particles and voids through lateral cross-section of shock tube. The camera records illuminated particles for analysis of particle concentration evolution and void structure analysis.](image2)
smoke or other gas tracers, we can perform LIF-based measurements of vertical motions in the shock tube.

Alternately, we can orient the optics such that the laser produces a horizontal light sheet through the shock tube, as we aim to measure the flux of glass particles as a function of time and distance from the initial bed height. As prior optical measurements performed by Heather Zunino have focused on the front of the cylindrical tube, this will allow us to see the entire cross-section of the flow and center of the tube, as long as the glass particle bed is sufficiently dilute to allow for penetration of the laser light sheet and optical access of the camera. These measurements will be advantageous in that voids in the interior of the tube may be visible and quantifiable. A schematic of this setup is shown in Figure 4.2. Note that this method is not meant to incorporate PIV analysis techniques, but it is possible we will be able to detect lateral motions through these experiments.

With the new timing system purchased from National Instruments, we will be able to synchronize the optical measurements with the pressure sensors. When the diaphragm bursts, the significant pressure drop above the diaphragm will be recorded via the PXIe-4492 card, and this signal will trigger the PCI-6602 card to initiate optical measurements by sending synchronized laser and camera pulses. The pressure sensors were not previously linked to optical measurements, so the new system provides an advantage in providing the relative timing between the diaphragm burst and onset of motion in the glass particle bed and in the surrounding gas.

In the initial PIV experiments Images were collected with the Phantom v641 high-speed camera. We synchronize the camera with the laser output by using a Digital Delay / Pulse Generator developed by Berkeley Nucleonics Corporation (BNC). The pulse generator has 8 available outputs, 4 of which control the laser flash lamps and Q-switches for each laser head, and 2 of which control the camera (one pulse per frame within a pair of images). At first, the PIV data collection operated continuously during experiments but wasn’t yet synchronized with the diaphragm burst or inception of the shock or rarefaction waves. The BNC signal generator has an available port to receive a trigger from an external source, and so we were able to incorporate a trigger generated by the pressure signal to be able to relate the timing of PIV images to the physical processes occurring within the shock tube. The initial images are shown below in Figure 4.3.

In exploring the timing of image collection in PIV, we found we are limited to a sampling frequency $f_s = 14.1$ Hz for maximum laser power, but we may be able to increase this frequency by diminishing the intensity levels of the laser heads. Within an image pair, we are able to use separation times ($\Delta t$) as low as approximately 5 microseconds. We are expecting to use values of $\Delta t$ around 50 $\mu$s to 100 $\mu$s according to our understanding of the speed of the waves from previous pressure measurements in Heather Zunino’s experiments.
While preparing the PIV experimental setup, we also worked to improve the shock tube itself, replacing parts of the nylon joints to ensure vertical alignment of the tubes and adequate sealing between sections. We also mounted a snorkel exhaust fume hood near the pump to collect stray particles that escape during the depressurization, as the silver-coated particles can be quite dangerous if inhaled.

4.1.5 Particle Image Velocimetry Triggered with Pressure Sensor Data

The team at ASU was able to trigger a particle image velocimetry (PIV) experiment with the pressure sensors, meaning we can measure the gas flow in the expansion region relative to the shock location in the shock tube. A brief set of PIV frames, along an image of the difference between the two frames is shown below in Figure 4.4.

4.1.6 Horizontal Void Cracks

As the dense particle bed expands, it has several interesting flow features. When the rarefaction wave impinges on the particle bed interface, the bed begins to swell and then is broken up by “horizontal void cracks.” These cracks then expand and break down further into smaller cells. By using some simple image processing techniques, the horizontal void cracks that appear in experiments with varying initial pressure ratios, particle sizes, and initial bed are plotted Figures...
4.5 – 4.7 below. The dark streaks are the cracks near the edge of the glass cylinder as they grow and move in time.

Figure 4.4: Shown are two PIV frames (left and middle) and the difference between the two frames (right), in which the upward motion is represented as traveling between dark pixels to bright pixels. The timing between the two PIV frames is 80 microseconds. The delay between the shock triggering the system to record and the first frame is 300 microseconds. The vertical field of view is approximately 3.5cm of gas and 0.5cm of the particle bed.

Figure 4.5: Bed composed of 212-297 micron particles. The independent variable in these plots is time (in seconds x \(10^{-4}\)) and the dependent variable is the axial coordinate along the length of shocktube (in pixels). Each of the six figures above is created by averaging the intensity across 25 pixels in a single image, plotted as a single column, and then the process is continued for nearly 350 frames. The first slice above contains information from the left wall of the shocktube to the first 25 pixels, the second slice is from 25 to 50 pixels, and the pattern continues until the right wall of the shocktube.
Figure 4.6: Bed composed of 150-212micron particles. This figure is created the same way as Figure 4.4.

Figure 4.7: Bed composed of 44 - 90micron particles. This figure is created the same way as Figures 4.5 and 4.6.
4.1.7 Pressure Sensor Data Examined with Differing Bed Heights

The pressure sensor data from 5 realizations of same experiment were ensemble averaged. Contained in each experimental data set, is the data from any redundant pressure sensors located at the same axial location along the shocktube. This yields four pressure traces. There are two above the diaphragm (1 pressure sensor at the very top and three 32cm above the diaphragm), showing the shock, and two below (2 pairs of pressure sensors), showing the rarefaction wave. See Figures 4.8-4.10.

Figure 4.8: Two experimental data sets with the same particle size range [44, 90] microns, but differing initial bed heights. Notice that the pressure traces located from above particle bed, seem unaffected. The rarefaction wave travels more slowly through the taller particle bed.
Figure 4.9: This plot is the constructed in the same way as Figure 4.8, however the particles in these experiments range between [150, 212] microns. The same effect is seen, although it is not as dramatic.

Figure 4.10: This plot is the constructed in the same way as Figures 4.8 and 4.9, however the particles in these experiments range between [212, 297] microns. Again, the same effect is seen, although it is even less dramatic.
4.1.8 Main Shocktube Setup

Additionally, new glass segments were ordered and installed on the shocktube. The new glass segments have minimal optical defects and have a much more uniform thickness and length than the previously installed glass segments. The new segments have been carefully measured and documented and that data has been sent to the team at UF.

With the new segments, the shocktube has been realigned so that the axis along the length (z-axis) is nearly perfectly perpendicular to the ground. There was an issue with some of the glass segments and connections not being perfectly straight, but that issue has been corrected and documented.

4.1.9 Fifth Year Plans

In 2018, the experimental team at ASU will complete the experiments. In addition, Heather Zunino is expected to graduate over the summer.

4.2 Eglin AFB Experiments

4.2.1 Goals and Motivation

The primary goal of the experiments conducted at Eglin Air Force Base is to provide validation quality data at the micro, meso, and macroscales. AFRL/RW has completed initial experiments at the micro- and meso-scales as described in this section, and the data have been given to the UF-CCMT modeling team for uncertainty analysis. The experiments include:

a. Microscale experiments with a no-particle, detonation-only, planar geometry in support of T3, uncertainties due to thermodynamic (EOS) and transport properties;
b. Mesoscale experiments with a planar particle bed in support of T5, compaction modeling uncertainty;
c. Microscale experiments with a planar particle cluster in support of T6/T7, uncertainty in finite Re, Ma and volume fraction dependent drag and heat transfer; and
d. Macroscale experiments with an annular particle bed in support of T0, validation of the demonstration problem.

4.2.2 Microscale Experiments

Twelve small-scale explosive experiments were performed at Eglin AFB in February 2015. These experiments are considered microscale in that a small number of particles are of interest. The test data include shock arrival times, high-speed video, x-ray images, and witness panel results. The twelve experiments are equally divided into gas-valve tests (compressed helium, tests 1-6) and explosively-driven test (tests 7-12). Table salt was substituted for tungsten particles as noted in the table in an attempt to visualize the gas flow patterns with the fine particles acting as tracers.

The microscale experiments were conducted at the Advanced Warhead Experimental Facility (AWEF) at Eglin AFB, FL. All tests utilize the same basic experimental setup. The pressure array
is a 3x3 grid of pressure probes, slightly offset from the test centerline with the middle probe removed due to the high likelihood of fragment impact. Four x-ray heads were used to increase the ability to track the particles’ location and velocity within the fireball.

Compressed helium provides an optically transparent shock wave and was used in the gas-valve experiments to refine the settings on the diagnostic equipment, particularly the cameras. During the second day of testing, the compressed helium driver was replaced by an explosive charge. For these experiments the explosive charge consisted of three stacked N5 pellets (each 0.5” in length and diameter) initiated by an RP-83 detonator.

Data from the experimental diagnostics are detailed below. For test 8, one of the delay generators malfunctioned, resulting in a loss of Simacon camera data. The delay generator was replaced in shot 9, but an improper arming sequence resulted in a loss of all data.

The pressure probes from the compressed helium tests show a relatively weak shock (≤1 psi) and a sharp pressure rise. Alternately, the pressure probes from the explosive tests show a much stronger shock (8-23 psi) with a complex pressure rise, exhibited in a non-noise signal, followed by an oscillation, then a sharp rise in signal.

High-speed images were recorded for the Phantom 6.11, Phantom Miro M310 and SIMACON. The Phantom 6.11 was used to capture the detonation products down range. The SIMACON camera was used to capture the blast wave at close range, but has a faster framing rate (but limited number of images) than the Phantom Miro M310. The Phantom Miro M310 was used to capture the blast wave at close range.

Four x-ray heads were mounted above the particle driver. The x-ray heads were triggered on a timing delay of 10-20 microseconds; the particle velocity is of primary interest. The large number and small size of the salt particles in test 10 and 11 precluded accurate velocity measurements.

Thin aluminum witness panels were used to determine the particles’ far-field impact locations in tests 7, 8, 9, and 12. The witness panels were placed 66” from the particle driver, and the center of each panel was determined using a bore sighting laser. No particles were located or recovered.
4.2.3 Mesoscale Experiments

Twenty-two mesoscale explosive experiments were performed at Eglin AFB in October-November 2015. The diagnostics and setup are the same as the microscale experiments. The 22 experiments are divided into gas-valve tests (tests 1-12) and explosively-driven tests (tests 13-22). The first tests were completed with a coarse steel powder. Tungsten powder was used for the remaining tests, where the best configuration of Tungsten required some experimentation, as seen in tests 8, 9, and 10. It was determined that the Tungsten powder was ejected most consistently when pressed in alcohol between sheets of tissue; tests 10-22 used this method.

4.2.4 Macroscale Experiments

The macroscale experiments were completed this past summer. Kyle Hughes (UF) was present at Eglin AFB during the testing to direct the testing so that satisfactory uncertainty quantification could be performed in future work. All six shots were successfully completed with accompanying video and pressure traces collected. Figure 4.11 shows the locations of the pressure probes for each of the experiments, as well as time sequence of a typical experiment. This data is currently being analyzed.

4.2.5 Summary

The micro and mesoscale experiments performed at Eglin AFB, FL, present the UF CCMT staff with a first-look at diagnostic and measurement techniques. The ultimate objective is to provide the UF-CCMT with high quality data for validation and verification of their turbulent multiphase models.
5. UB Team

5.1 Summary of Achievements

1. Validation and uncertainty quantification (UQ) of the mesoscale shock tube simulation
   a. Wrapping up validation and UQ of the mesoscale shocktube simulation
   b. Found that UQ was very useful for identifying modeling issues in the simulation under development
   c. Found that experimental data is often obtained by elaborate measurement processing, which is one of major uncertainty sources
   d. Achieved the grid convergence of the hydrodynamic force in a Euler-Lagrange simulation for a single particle by using a finite-size particle model
   e. Quantified the discretization error for quasi-steady, pressure gradient and added mass force for a single particle
   f. Noise characterization and reduction in the prediction metric of the macroscale cylindrical detonation model

2. UQ of Eglin experiments
   a. Lessons learned and uncertainty quantification results obtained from investigation of past microscale experiments by an independent investigator, labeled forensic uncertainty quantification, is being currently being compiled into a journal article
   b. Mesoscale gas gun experimental design were significantly improved to meet simulation capabilities in initial testing. However, equipment failure prevented completion of the tests
   c. Successful firing of the macroscale tests were performed. Uncertainty quantification personnel were on-site to ensure gathering of all necessary uncertainty quantification data and documentation of the testing process

3. Multi-fidelity surrogate modeling for application/architecture co-design
   a. Designed the experiments for comparing CMT-nek (Jason Hackl), CMT-bone (Tania Banerjee), CMT-bone BE (Aravind Neelakantan) and BE emulation (Aravind Neelakantan)
   b. Proposed a multi-fidelity surrogate to fit a few CMT-nek runs together with BE emulations for improved accuracy.
   c. Applied and validated the multi-fidelity surrogate for extrapolation towards large-scale runs.

5.2 Overview

The primary objective of the UB team is to estimate the error in compressible multiphase turbulence (CMT) models to assess the prediction capability of the hero simulation based on the model. Often model error estimation is suffered by large uncertainty. Thus, uncertainty reduction (UR) was emphasized and UQ was utilized to prepare UR. For systematic UR, the effects of the uncertainty sources on the total uncertainty were categorized and UR priority was made for reducible epistemic uncertainties.
For the Sandia shocktube simulation, about 70% of the initial uncertainty has been reduced. It was found that UQ was very useful for identifying invalidities of the simulation under development and experiment measurement process could be a large uncertainty source.

Validation and UQ of simulation and experiment coupling efforts have been being made based on the UQ and validation framework that was developed for the shock tube simulation. The Eglin macroscale experiments were carried out in 2017, UQ of the experiments are being carried out. With limited resource, we planned the experiments for minimizing uncertainty. For example, to control the uncertainty due to imperfection of particle casing, we decided to fabricate the casing with imperfections designed to control the consequence of the imperfection rather than let it play as a random effect. It was discovered that the influence of casing was little.

A secondary objective is to help other teams quantify and reduce the uncertainty in their applications. With the exascale team, validation and UQ of CMT-bone and CMT-bone-BE are being carried out.

5.3 Validation, Uncertainty Quantification and Uncertainty Budget of Mesoscale Sandia Shock Tube Simulation

Two largest epistemic uncertainties in the Sandia shocktube experiment were: 1) uncertainty due to gap and 2) uncertainty in the initial volume fraction. They contributed more than 90% of the total uncertainty, which made the estimation of the error uncertain.

Figure 5.1 (a) shows the estimate error in the downstream front position at 700 µs. The error is about -2% to 10% in terms of percentage. The prediction capability of the simulation was uncertain. New experiments were conducted without gaps. The new experiments revealed that the effect of gap was little. The measured prediction metrics from the experiments with gap and without gap were almost identical. After incorporating the data from the new experiments, the uncertainty due to gap was significantly reduced. Figure 5.2 (b) shows two error bars before (black) and after (red) the UR.

The second largest epistemic uncertainty is the uncertainty in the volume fraction. The uncertainty was also an uncertainty from the process to obtain the volume fraction from the X-ray image of the particle curtain. It was simply quantified with conservatism to compensate the simplicity. A new rigorous uncertainty study reduced the uncertainty more than 50%. The uncertainty reduction has been reflected in the model error estimation in Fig. 5.1 (c). The red error bar is the most recent error bar. It shows that the initial uncertainty in the error estimation in Fig. 5.1 (a) has been significantly reduced.
After the second round UR, it was concluded that the remaining uncertainty was no longer an obstacle and no more UR was required. The most suspected error sources were the error in the particle force model and the error in the volume fraction profile due to the use of top-hat volume fraction profile.

The particle force model calculates force on a particle by summing five force components: quasi-static, added mass, inviscid, pressure gradient and particle interaction forces. Samaun Nili is particularly devoted to rank the models to calculate force components of particle force in terms of their contributions to the model error. The simulation is based on the Euler-Lagrange equation that cause difficulty for achieving convergence of the simulation. He is also working on achieving convergence using the method of finite size particles based on the generalized Faxen theorem.
5.4 Validation, Uncertainty Quantification and Uncertainty Budget of Mesoscale ASU Shock Tube Simulation

The validation and UQ of the ASU simulation is currently being engaged. Figure 5.2 shows a schematic figure of the experimental configuration. The vertical shocktube is composed of four sections. A vacuum is created in Section A and B and those sections are separated from Section C and D in atmospheric pressure by a diaphragm. When the diaphragm bursts, an expansion fan is created with a shock and it goes through the particle bed in Section D. The particle bed expands and the height of particle bed increases. The particle bed height is the prediction metric.

As a preliminary UQ study, five simulation runs were made with randomly generated particle positions and were compared to the experimental measurements. Figure 5.2 shows, a comparison

Figure 5.2. A schematic figure of the experimental configuration and a comparison between the simulations and experiments with random particle positions

Figure 5.3. UQ of a particle bed using a mockup

(a) Regions where volume fractions were measured
(b) Volume fraction at different slices (the range of 500 to 1500 slice is equivalent to 0 to 360 degrees)
between the four experiments (dashed lines) and five simulations runs using CMT-nek in terms of the particle bed height. The simulation runs indicated that the influence of the random effects on the particle bed height was little. It was because the randomness in individual particle positions were averaged out by the particles more than a million.

UQ of the experimental results is being carried out. Table 5.1 shows uncertainty sources in the experiments. We found that the uncertainty in the parameters were little except the pressure ratio. The uncertainty in the pressure ration is because of the uncertainty in the pressure in the vacuum. The uncertainty was recognized by finding inconsistency between the pressure ration needed to generate the shock speed and the measured pressure ratio. Since the pressure of Section A and B were measured by a static pressure sensor planted on top of Section A and there was a leakage around the diaphragm, the pressure in the vacuum was questionable. UQ of the pressure ration is being carried out.

One of the efforts is to measure the volume fraction of a particle bed. A mock-up particle bed was made and the volume fraction of the bed was measured with a CT scanning image of the bed. Figure 5.3 (a) shows five regions where volume fractions were measured. Figure 5.3 (b) shows measured volume fraction in the regions for different slices.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Quantity</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume of particle bed</td>
<td>383.5-385.6 cm³</td>
<td>CT scanner</td>
</tr>
<tr>
<td>Initial volume fraction</td>
<td>59-65%</td>
<td>CT scanner</td>
</tr>
<tr>
<td>Initial mean volume fraction</td>
<td>60-61%</td>
<td>$\phi_{mean} = \frac{m}{hA} * \frac{1}{\rho_{glass}}$ (Justin)</td>
</tr>
<tr>
<td>Particle diameter (before)</td>
<td>67±8 μm</td>
<td>SEM</td>
</tr>
<tr>
<td>Particle diameter (after)</td>
<td>68±22 μm</td>
<td>SEM</td>
</tr>
<tr>
<td>Particle density</td>
<td>2.44±0.006 g/cm³</td>
<td>Pycnometer</td>
</tr>
<tr>
<td>Driver section pressure</td>
<td>100.68 kPa</td>
<td>Static sensor</td>
</tr>
<tr>
<td>Driven section pressure</td>
<td>5.5±? kPa</td>
<td>Static sensor on the ceiling</td>
</tr>
<tr>
<td>Tube geometry</td>
<td>Negligible</td>
<td></td>
</tr>
</tbody>
</table>

5.5 Uncertainty Quantification (UQ) of Eglin Experiments

5.5.1 Uncertainty quantification for Eglin microscale experiments.

A promising new approach to performing uncertainty quantification (UQ) is being developed for past experiments that involves the adoption of the perspective of a crime scene investigator. Like a crime scene investigator, the UQ personnel collects and documents “evidence” about the experiment to reconstruct an accurate picture of the experiment. In addition, after-the-fact characterization of the measured inputs may be performed in the laboratory to confirm details, similar to a forensics laboratory. In addition, it is postulated that there is a distinct advantage in have a third party beside the simulationist and the experimentalist perform the investigation. The relative independence of the investigator allows for a critical evaluation of discrepancies between
experiments and simulations. A journal article is currently being drafted, to be submitted to AIAA Journal, exploring the concept of forensic UQ with application to the microscale experiments.

5.5.2 Uncertainty quantification for Eglin mesoscale experiments.

Kyle Hughes (UF) assisted Mike Jenkins (AFRL) as they attempted to complete the gas gun experiments. However, the mesoscale experiments were unable to be completed this summer due to equipment issues. The laser synchronizer was faulty and had to be sent to the manufacturer for repairs. The tests will be resumed once the PIV setup is returned to operational status and time is available in the testing chamber.

One of the challenges for the current setup is the fragmentation of the sabot, Figure 5.4. Large pieces of sabot can be seen following immediately after the particles and is a significant source of uncertainty. The velocity of the sabot is being reduced to eliminate the fragmentation while still maintaining large particle velocities.

![Figure 5.4. Test firing of a sabot carrying 1.0 grams of tungsten powder. The sabot impacts the stripper plate at the 0 µs mark. The dark cloud is the ejection of the tungsten powder and then the white sabot fragments can be seen quickly overtaking the powder.](image)

5.5.3 Uncertainty quantification for Eglin macroscale experiments

The macroscale experiments were completed this past summer. Kyle Hughes (UF) was present at Eglin AFB during the testing to direct the testing so that satisfactory uncertainty quantification could be performed in future work. All six shots were successfully completed with accompanying video and pressure traces collected. Still images are presented in Figure 5.5 to demonstrate the
surprising difference in character between the steel and tungsten shots despite similar mass to charge ratios.

During the planning stages there were two major concerns raised:

1) Will the casing affect the results, especially with regards to the formation of instabilities?
2) Will instabilities form at all?

Figure 5.5 shows the formation of the instabilities 1.67 ms after detonation as the particle liner emerges from behind the soot cloud, answering the second concern. As for the first concern, examination of the pressure traces in Figure 5.6, shows the remarkable agreement between the three repeated steel liner shots. The notching does not appear to affect the shock time of arrival significantly. Qualitative examination of the high-speed videos shows no obvious effect on the instabilities. Future work will examine the data for effects on the instabilities formed.
Figure 5.7 shows the image analysis results for each of the particle liner shots along with the shock TOA results from the pressure probes. Camera 3 allows tracking of the shock normal to the ground. Pressure probes referenced here are those lying on the centerline of the cylindrical charge (normal to the long axis and along the ground). The redundant instrumentation provided several benefits. First, the camera provides an excellent complement to the pressure probe data, especially by providing data at many more points than the four positions offered by the pressure probes if with slightly greater variance.

Second, the redundant instrumentation can start to rule out compounding factors such as ground effects. The close agreement between the pressure probes and camera 3 for the steel shot indicate that ground effects do not largely affect the shock TOA for the steel liner configuration. The tungsten shot exhibits the opposite behavior of what one expects if observing ground effects retarding the shock traveling over the pressure probes. The image analysis shows a slower shock traveling away normal to the ground. The pressure transducers were oversized for the first particle liner shot (tungsten) and better sized in subsequent shots (steel). It is postulated that the pressure probes are sluggish to respond for the tungsten shot since they are responding to relatively small pressure wave with respect to their full dynamic range.

Figure 5.6. Shock time of arrivals obtained from pressure records for the three steel shots. The first two shots contained notching while the third was unperturbed.
5.6 Error Quantification of Particle Force Mode for 1D Shock-Particle Interaction via Model improvement

This study presents an approach to model the interaction between particles and fluid for finite-size particles that permits convergence. Here we use the generalized Faxen form to compute the force on a particle and examine the limit of finite-size particle to compare the results against traditional point particle method. This is done by apportioning the different force components on the particle to fluid cells based on the fraction of particle volume in the cell (e.g., pressure gradient and inviscid unsteady force) or the fraction of particle surface area in the cell (e.g., quasi-steady force). The application is to a one-dimensional model of shock propagation through a particle-laden field of
spherical particles at moderate volume, where the convergence is achieved for a well-formulated force model and back coupling for finite-size particles.

So far for a single stationary and moving particle, the convergence is achieved for a well-formulated force model and back coupling for the finite-size particle. We illustrated the force on the particle and its projection on the grid cells for finite-size particle model versus the traditional point particle model in Fig. 5.8. The particle is about 18 times larger than the mesh size on this example. In Fig 5.8, \( t^* \) is the normalized time which equals to the ratio of the actual time \( t \) to the time it takes for the shock to travel over the particle diameter \( D_p \). \( U_S \) denotes the speed of shock. Figure 5.8 (a) compares the force for finite-size particle against traditional point particle method when a shock wave passes through the particle where the total force is equal to the sum of the quasi-steady, pressure gradient and added mass forces. In point particle model, we use linear interpolation to estimate the gas properties and the forces at the center of the particle. As Fig. 5.8 (a) shows, the time evolution of point particle has a very sharp peak because it senses the shock much later than when the particle is being touched by the shock. Whereas the correct force (blue line) is being calculated by finite-size particle model where the forces, are evaluated not based on the value of the fluid properties at the center of the particle, but we calculated them based on surface and volume average gas properties. Hence, the particle sees the forces as soon as it is being touched by the shock. Also, in point particle approach, the force will be put back only to the two or three grids around where the particle center lies. But in finite-size particle approach, we already know the contribution of forces from each of these segments. So, each of these segments can feedback to its neighboring cells using discrete delta function. Figure 5.8 (b) illustrates the projected particle force on the grids for finite-size particle approximation. Sines the particle occupied by 18 cells, when we use discrete delta function, it put puts back the force to the 20 grids around the particle.

In Fig. 5.9, the particle is about 4 times larger than mesh size, and the particle to fluid density ratio is about 10. We plotted the force projected onto 11 different grid points in Fig. 5.9. In this scenario, only 4 to 5 cells are resolving the particle, and we monitor the force on all these cells. Figure 5.9 (a) showed the conventional point particle approach where we used linear interpolation to put the force back to the cells; each grid force starts to increase as the particle approaches to the cell center and then starts decreasing as it gets further. So, there is a lot of ups and downs created these sharp discontinuities. Whereas in the finite-size point particle approach shown in Fig. 5.9 (b), each cell can smoothly see each contribution of feedback force which slowly increases, reaching the peak and it goes down. Here we have a two-way coupled simulation where we think the finite-size particle approximation will provide us a converged solution by reducing these sharp changes in the force.
For the next step, we are planning to apply finite-size particle model for a one-dimensional model of shock propagation through a particle-laden field at moderate volume fraction. Also, compare the results with Direct Numerical Simulation (DNS) will be used to check if the approach also improves accuracy compared to the point particle model.

5.7 Multi-Fidelity Surrogate Modeling for Application/Architecture Co-design

UB team has been working together with Nek team, CS team and Exascale team to quantify the similarity and difference between run times of CMT codes at large-scale runs. The study is based on the execution time of a typical no-particle test (run on Vulcan) using up to 34 million \((131072 \times 256)\) elements and 311 billion \((21^3 \times 131072 \times 256)\) computational grid points. We have achieved 22 runs from CMT-nek, 67 runs from CMT-bone, 125 runs from CMT-bone BE and BE emulation as shown in Fig. 5.10 (a). We explored multi-fidelity surrogate (MFS) approach to combines data samples of low-fidelity (LF) models (BE, CMT-bone BE, CMT-bone) with a few samples of a high-fidelity (HF) model (CMT-nek).

We investigated the Multi-fidelity model based on rich BE simulations (LF) and a few CMT-nek (HF parent app) data points to predict the performance of CMT-nek (HF). 10 runs (out of 22) were selected randomly and fixed as the validation runs. Different combinations of training and test set were generated from the remaining 12 runs. The MFS predictions for CMT-nek runs with increasing number of samples were summarized in Fig. 5.10 (b). The relative root-mean-square errors were less than 10 % with more than 9 CMT-nek samples and ended with 4.49%. The MFS was much more accurate than either the LF model or the linear fit to HF data, demonstrating its promise to compensate the difference between high-fidelity and low-fidelity models.
Our ultimate goal is to predict the performance for exascale computation platform which is essentially long-range extrapolation far from the training samples. We developed MFS for long-range extrapolation using 7 (out of 22) runs of CMT-nek (high-fidelity), at 256 and 2k processors. The predictions of MFS were evaluated towards 16k (11 runs) & 128k (4 runs) processors. The average error at the 15 large-scale validations runs was 6% and the maximum error was 11%. Even for long-range extrapolation, the MFS proved to be promising for the selected benchmark study.
6. CMT-nek Code Development Team

6.1 Overview

The primary code utilized at Florida is CMT-nek, an extension of the award winning, million-core scalable, spectral element code Nek5000. CMT-nek solves the three-dimensional time-dependent compressible Navier-Stokes equations on a locally structured, globally unstructured grid. The grid is made of up hexahedral elements in 3D and quadrilaterals in 2D. The spatial discretization is carried out using spectral elements, employing a discontinuous Galerkin method. The inviscid fluxes are approximated by upwind schemes to allow for capturing of shock waves and contact discontinuities. The viscous fluxes are approximated using the Baumann-Oden method. The code will also employ an entropy-viscosity method for stabilization. Flows are computed with the third-order accurate Runge-Kutta method for time marching. Typical production run sizes that our team has run over the past year ranges from 4000 cores to 100k cores.

With a complete implementation of the smoothing and tuning procedures in the entropy viscosity method (EVM, Guermond, Pasquetti & Popov (2011) J. Comp. Phys. 230:4248-4268) adapted to the discontinuous Galerkin spectral element method (DGSEM), 2017 was devoted to convergence studies for V&V and publication as well as identifying parameters and practices for CMT-nek as a more mature shock-capturing solver of the Euler equations of gas dynamics. A mollification approach to point-particle forces and the discrete element method (DEM) for collisions enabled full four-way coupling between particles and between particles and gas. CMT-nek began production use in CCMT with runs for simulating the Arizona State University shock tube. Microscale flows received significant attention in 2017, wrapping up the simulations of potential flow over an array of spheres and producing CMT-nek’s first simulations of shock waves interacting with resolved spheres. This focus is shifting as Year 4 ends, with the start of a full migration of CCMT to CMT-nek for the macroscale demonstration problem for blast waves interacting with low volume fractions of particles.

6.2 Artificial viscosity and convergence studies

Artificial viscosity in CMT-nek at the end of 2016 was originally piecewise constant within elements, and property

Figure 6.1. CMT-nek solution at polynomial order \( p=8 \) of density in a collision between two shock waves in a polytropic gas at \( t=0.035 \). The least-squares fit to the \( L_1 \) error in element-averaged density gives an empirical order of convergence of 0.996 in mesh spacing \( h \).

The summer was devoted to convergence studies of canonical flows with shock waves that appear in a paper (Hackl, Shringarpure, Koneru, Delchini and Balachandar) under review at Computers & Fluids detailing the full formulation of the Eulerian solver in CMT-nek. Figure 6.1 shows CMT-nek’s first-order convergence to an exact Riemann solver’s solution of the collision between two shock waves. The empirical rate of $h^{0.996}$ is comparable to the $p$-convergence rate at $h=0.01$, and such computations for the Sod (1978) shock tube meet or exceed the empirical convergence rates of other shock-capturing discontinuous Galerkin schemes reported in the literature over one decade in mesh spacing and polynomial orders as high as 16. These convergence studies also produced values for tunable coefficients for the entropy viscosity method used in microscale simulations and the transition to macroscale multiphase work.

6.3 Microscale simulations

In 2017, we simulated expansion waves propagating into particle beds at higher volume fractions in order to validate the latest drag models for spheres in a compressible flow. Previously, a particle bed of 3% volume fraction was simulated using CMT-nek. We extended that to particle beds of volume fractions 10% and 15%. The drag model predicted the actual drag forces very well up to the time waves reflecting off downstream particles were felt upstream (Figure 6.2). Further, by modeling the particle bed as a sudden area change, we were able to predict the final pressure, temperature and Mach number using an isentropic flow analysis.

Drag from shock waves propagating over spheres are an important building block of modeling forces on particles in multiphase flows, and in preparation of wider CMT-nek use by the
Microscale group shock-sphere interactions at two shock Mach numbers (1.22 and 3) were also conducted on finely resolved fully three-dimensional meshes. Their drag time series in Figure 6.3 are validated against axisymmetric simulations in RocFlu (Mehta et al. 2016). CMT-nek predicts the Mach number dependence of the diffracted shock structure, resolving the contact discontinuity.

Figure 6.3. Drag time series from CMT-nek simulations of shock waves interacting with spheres at two shock Mach numbers. Time normalized by traversal time scale based on shock speed $V_s$ and sphere diameter $D$. Results from legacy code RocFlu appear in Mehta et al (2016) J. Appl. Phys. 119(10):104901.

Figure 6.4. Mach number contours overlaid on a pseudocolor plots of normalized pressure $p/p_\infty$ (top half) and artificial viscosity coefficient ("s1," bottom half) at $tV_s/D = 1$ in the CMT-nek simulation of a Mach 1.22 shock interacting with a rigid sphere.

Figure 6.5. Density pseudocolor and temperature contours on y- and z-planes at $tV_s/D=1$ in the CMT-nek simulation of a Mach 3 shock interacting with a rigid sphere. The triple point structure resembles the Schlieren observations of Bryson & Gross (1960) J. Fluid Mech. 10(1):1-16.
at the (revolved) triple point between the Mach 3 incident shock and its reflected and diffracted companions (Figure 6.5) that is not present at a shock Mach number of 1.22 (Figure 6.4). These simulations used a 1M-element mesh generated in GridPro, a polynomial order of 4, and took 366K CPU hours on Vulcan and 445K CPU hours on Quartz at LLNL.

6.4 Lagrangian point-particle modeling and algorithms

In 2017 the Lagrangian point-particle suite in CMT-nek transformed substantially, taking the leap into four-way coupling and improving its massively parallel processing capabilities. Evaluating a Gaussian mollification function with sub-element standard deviation centered at a given particle position smoothly back-couple’s that particle’s forces on the gas (with or without superparticle loading) to surrounding grid points, resolving Eulerian distributions of particle volume fraction for the first time. The soft-sphere discrete element method (DEM, Cundall & Strack (1979)) enables CMT-nek to track collisions and arrange Lagrangian point particles into physically meaningful packs when they are close to one another. Back coupling from particles within some distance of an element’s faces via a “ghost particle method” (illustrated in Figure 6.6) more gently integrates particle forces into the discontinuous Galerkin framework than an entirely element-local approach would allow and avoids collision detection errors at faces.

Extensive calibration and validation for publication lies ahead, but the Lagrangian capabilities of CMT-nek saw extensive application in 2017 to the Arizona State University shock tube in which a particle bed initially placed behind the driver diaphragm are evacuated by the shock tube’s expansion fan. Figure 6.7 shows volume fraction from a later time in a 20K-element simulation using 8192 cores on Vulcan at LLNL.

6.5 Fifth Year Plans

Large (73M particles, 1.2M-element) simulations of the ASU case are in preparation along with manuscripts for publishing the fully coupled multiphase flow solver that CMT-nek now implements. Migrating macroscale group attention to conducting simulations (including hero runs) of the demonstration problem is the central priority in the first quarter of 2018, with a goal of completing a larger version of CCMT’s first hero run (perfect gas at detonation state) by the end
of March, 2018. Second quarter will be devoted to adding a detonation-product phase and migrating RocFlu’s equation-of-state surrogate models to CMT-nek. Our ambitious goal is to reproduce full packings of rigid spheres (60% solid volume fraction) in blast problems by the end of the year.

Figure 6.7. Particle volume fraction $\phi_p$ 0.015s into a simulation of the interaction between an expansion wave and a bed of particles at the conditions of the ASU shock tube.
7. CS Team

7.1 Overview
The research effort of our CS team has been to optimize CMT-nek in terms of performance, power and energy consumption. Towards that end, we have been working on load balancing CMT-nek. We have also developed a dynamic programming approach for energy and power optimization on a DVFS platform.

7.2 Load balancing CMT-nek
The steps and strategy adopted for load balancing CMT-nek are described below.

7.2.1 Determining Computational Load on a Spectral Element
Computational load on an element is quantified as the number of particles present inside the element plus a baseline load for fluid computation. The latter is about the same for any element since fluid computations involve solving fluid properties at each grid point in an element and all hexahedral spectral elements have the same number of grid points. In fact, the load on each element due to fluid computations is \( O(N^4) \), where \( N \) is the number of grid points along one direction. To ensure that particle load does not dominate fluid load and vice versa, we represent fluid load as a constant defined as the ratio of the average time it takes to process a single element to the average time it takes to process a single particle by running the application prior to enabling load balancing. This constant is dependent on the platform and on problem parameters such as grid size. Once the computational load for each element is determined, the next step is to repartition the 1-D array of spectral elements which is described next.

7.2.2 Domain Repartitioning Strategies
We developed a centralized, a distributed and a hybrid algorithm for repartitioning the 1-D array of spectral elements. The 1-D array of spectral elements is formed by executing a recursive spectral bisection algorithm on the fluid domain.

Centralized. The main theme of the centralized algorithm is that all the processors send their total computational load to processor P0. Processor P0 then computes the prefix sum of the load on elements ordered according to their global IDs, partitions the prefix sum array, uses the partitioning to create a new element \( \rightarrow \) processor map \( M' \), and finally, broadcasts \( M' \) to all processors. The algorithm and an example may be found in our submitted paper [3].

Distributed. The centralized load balancing has processor P0 in the critical path that determines how fast the load balancing may complete. In the distributed version, we remove this bottleneck and let each processor collaborate to have a local copy of the prefix sum of the load. After that, each processor calculates a local element \( \rightarrow \) processor map. The processors share their local maps which each processor composes to form a global element \( \rightarrow \) processor map. As a last step, each processor adjusts the mapping to guarantee that the number of elements assigned to a processor
does not exceed a maximum bound defined by the user by setting a variable “lelt”. The distributed load-balancing scheme and an example may be found in our submitted paper [3].

Hybrid. The hybrid load-balancing algorithm is a combination of the centralized and distributed algorithms. First, the processors collaboratively create the local copy of the prefix sum and each processor calculates a local element→processor map. Then, each processor sends the local map to processor P0. P0 then aggregates the data to create the global map. P0 also adjusts the mapping to guarantee that the number of elements assigned to a processor does not exceed “lelt”. Finally, processor P0 broadcasts the element→processor map, M’, to all processors. The hybrid load-balancing scheme and an example may be found in our submitted paper [3].

7.2.3 Distributing Elements and Particles
Here we will describe the details of how data is transferred between processors to achieve load balance. There are many data structures in CMT-nek that contain element and particle information. Some of these data structures store static data while others store dynamic data. Static data include information such as the x, y and z coordinates of each element, curvature on the curved faces, and the boundary conditions. Dynamic data includes fluid and particle properties that change during simulation. We follow two primary strategies for information transfer.

7.2.4 Transferring Data
Arrays storing the conserved variables such as mass, energy, and the three components of momentum are transferred. The transfer process consists of packing the arrays to be transferred, transferring the packed array, and finally unpacking the arrays received. We use the underlying crystal router in Nek5000 for transferring the packed data.

7.2.5 Reinitializing Data
The data structures which store static data are reinitialized. For this step, we initiate calls to existing Nek5000 data initialization routines. This needed careful analysis since some routines store the status of calls using static variables local to the scope of the routines. We analyzed and updated such routines to reset the state of those variables when load balancing is done. After load balancing is complete, a processor will start the next time step, which involves the computation of field variables for all elements, including the new ones that were received. For all tests, we have diligently verified that the results of simulation have the same accuracy as the original CMT-nek.

7.2.6 Automatic Load Balancing
As particles move during simulation, the original element→processor mapping becomes suboptimal since the particle-heavy elements with large computational load start getting lighter on particles as the particles move apart. This motivates the need for a dynamic load-balancing scheme, where load balancing may be triggered during an ongoing simulation process. There are two main strategies for triggering load balancing during simulation in CMT-nek. The first is performing load
balancing at specific intervals where the intervals are specified by the user. The second strategy is adaptive load balancing, where no input is necessary from the user.

**Fixed Step Load Balancing.** This type of load balancing requires user input. For example, a load-balancing step may be triggered after every k time steps, where k is specified by the user. To set a reasonable value for k, the user should be aware of the simulation details, such as the problem size, particle speed, duration of a time step, and so on. The user may also run the simulation for a certain time without load balancing to estimate k.

**Adaptive Load Balancing.** This type of load balancing is performed automatically by the program, with no input from the user. The details of the adaptive load-balance algorithm are presented in our paper [3]. Before the adaptive load balancing strategy is applied, there is one compulsory load-balancing step which happens at the beginning after particles are initialized and placed and before the start of simulation. After that a load-balancing step happens whenever performance degrades by a certain threshold. For subsequent load balancing steps, the application captures the cost of load balancing and the slope of performance degradation. Using these, we can get the theoretical load-balancing interval. When this interval is reached, the load-balancing algorithm is called. However, since the slope may be changing, we add another criterion which is: when the cost of load balancing is covered by the cost caused by the increasing time-per-time-step, a load balancing step is called. After that, each time a load balancing is performed, the cost of load balancing and the slope are updated, and the new theoretical load-balancing interval is calculated. Thus, load balancing happens when either one of the two criteria is satisfied.

### 7.2.7 Experimental Results
#### 7.2.7.1 Problem Description
To test the load-balancing algorithm, we used a test case that has been devised to mimic some of the key features of particle-laden, explosively driven flows that the load-balancing algorithm proposes to overcome. The test deals with expansion fans in one dimension which are simple compressible flows. The problem domain is a rectangular prism that extends from 0 to 0.0802 in the y and z directions and from −2.208 to 6.0 in the x direction. Note that the units in this case are non-dimensional. The particles are assigned between −1.0 and −0.5 in x direction, where the difference between the left (x = −1.0) and right (x = −0.5) boundaries determines the initial volume fraction of particles. The left boundary is often adjusted to obtain a different initial volume fraction.

#### 7.2.7.2 Platforms
Quartz is an Intel Xeon platform in Lawrence Livermore National Laboratory. Quartz has a total of 2,688 nodes, each having 36 cores. Each node is a dual socket Intel 18-core Xeon E5-2695 v4 processor (code name: Broadwell) operating at 2.1 GHz clock speed. The memory on each node is 128 GB. Quartz uses Omni-Path switch to interconnect with the parallel file system.
Vulcan is an IBM BG/Q platform in the Lawrence Livermore National Laboratory. Vulcan has
24,576 nodes with 16 cores per node, for a total of 393,216 cores. Each
core is an IBM PowerPC A2 processor operating at 1.6 GHz clock speed. The memory on each
node is 16GB. Vulcan uses a 5-D torus network to interconnect with the parallel file system.

7.2.7.3 Experiments on Quartz

Figure 7.1 shows the overhead of the centralized, distributed and hybrid load-balancing algorithms
on Quartz. It is a weak scaling with 4 elements per MPI rank and about 343 particles on an average
per element. The variable \( \text{elt} \) was set to 16. Each spectral element consists of \( 5 \times 5 \times 5 \) grid points.
The overhead includes time taken for each of the following steps: 1) remapping elements to
processors; 2) packing, sending, and unpacking received elements and particles; and 3)
reinitialization of data structures that are used in computation. The horizontal axis represents the
number of MPI ranks while the vertical axis represents the time in seconds taken to load balance
the application.

![Overhead of load balancing algorithms on Quartz](image_url)

Figure 7.1. On Quartz, total overhead for a load balancing step for centralized, distributed and hybrid
algorithms. It is a weak scaling with 4 elements per MPI rank, \( 5 \times 5 \times 5 \) grid points per element, and
about 343 particles per element. The actual overhead expressed as number of time steps for 65, 520 MPI
ranks was 1.94 for the centralized, 3.35 for the distributed, and 1.82 for the hybrid algorithm.

The overhead incurred by a load-balancing step increases with the number of MPI ranks. Ideally,
the distributed algorithm should take less time than the centralized algorithm with increasing MPI
ranks since there is no processor P0 bottleneck in it. However, on Quartz the centralized algorithm
is faster due to a higher ratio of communication-time to computation-time on the system and the
distributed algorithm is rich in communication especially in MPI\_ALLGATHERV. The hybrid
algorithm, eliminates calls to MPI\_ALLGATHERV, as well as, the part in the centralized
algorithm where all processors send their element loads to P0. As we can see from Figure 1.1, the
hybrid algorithm was the fastest. The actual overhead for 65, 520 MPI ranks for centralized,
distributed and hybrid was 0.33, 0.57 and 0.31 seconds, respectively. Compared to the time per
time step which was 0.17 seconds, the overhead expressed as number of time steps was 1.94 for the centralized, 3.35 for the distributed, and 1.82 for the hybrid algorithm. This makes dynamic load balancing practical for a large class of simulations. For these experiments, the total number of time steps was 100, and load-balancing took place every 10 steps. Thus, we found that the overhead for load balancing is low and scales very well with the number of processors.

The load-balanced and non-load-balanced (original) codes were run on 67, 206 MPI ranks on Quartz, that is 1, 867 nodes with 36 cores per node. The grid size per element was $5 \times 5 \times 5$ and the total number of elements was 900, 000. The variable $l_{el}$ was set to 120 elements. The total number of particles was $1.125 \times 10^9$, obtained as 1250 particles per element on an average. Initially, the percent of elements that have particles is 6.1% of the total number of elements.

Figure 7.2 compares a trace of the CPU time taken per simulation time step for load-balanced versus the original. Adaptive hybrid load balancing was used in this example. The average time per time step for the original and the load balanced versions were 9.92 and 0.995 seconds, respectively. Thus, we gained an overall speedup of 9.97 using load balancing algorithm. During the duration of simulation, apart from the compulsory load balancing that happens before simulation time step 1, CMT-nek begins load balanced at 4, 077 time steps and after that, it will automatically load balance by itself (the small blue dot there represents the time taken by next step after load balance). Original version did not finish in 2.2 hours.

Figure 7.2. Performance comparison between load-balanced and original versions of CMT-nek on Quartz. They were run on 67,206 MPI ranks, that is 1,867 nodes with 36 cores per node. Adaptive hybrid load balancing was used. The average time per time step taken by the original version and the load balanced version were 9.92 and 0.995 seconds, respectively, giving us an overall speed-up factor of 9.97. Original version did not finish in 2.2 hours.
7.2.7.4 Experiments on Vulcan

We now evaluate the load-balancing algorithms on Vulcan. Figure 1.3 shows the total overhead for a load-balancing step using the centralized, distributed and hybrid load-balancing algorithms. It is a weak scaling study so problem size increases proportionally to the number of MPI ranks, which is 2 elements per MPI rank and 343 particles per element on average. As we can see from Figure 7.3, the load-balancing overhead increases with an increasing number of total MPI ranks. Especially, distributed algorithm was faster than the centralized and hybrid algorithms. That is because of a lower ratio of communication-time to computation-time on this platform. The actual overhead for 393,216 MPI ranks for centralized, distributed and hybrid algorithm was 1.00, 0.77 and 0.84 seconds respectively. Compared to the time per time step which was 0.33 seconds, the overhead expressed as the number of time steps was 3.03 time steps for the centralized, 2.33 for the distributed, and 2.55 for the hybrid algorithm. The variable lelt, which is the maximum number of elements on an MPI rank, was set to 8 for these overhead runs. Of these experiments, the total number of time step was 100 and load balancing took place every 10 steps. Again, we can see from these results that the overhead for load balancing is low and scales very well with the number of processors.

The load-balanced and original codes were run on 65,536 MPI ranks, that is 16, 384 nodes with 4 cores per node. The grid size per element was 5x 5x 5 and the total number of elements was 900, 000. The total number of particles was 1.125 x 10⁹ obtained as 1250 particles per element on average. Initially, the percent of elements that have particles was 6.1% of the total number of elements. The variable lelt was set to 140 for the load-balanced version. Figure 7.4 shows the

![Figure 7.3](image-url)

Figure 7.3. On Vulcan, total overhead for a load-balancing step for centralized, distributed and hybrid algorithms. It is a weak scaling with 2 elements per MPI rank, 5 x 5 x 5 grid points per element, and about 343 particles per element. The actual overhead expressed as the number of time steps for 393, 216 MPI ranks was 3.03 for the centralized, 2.33 for the distributed, and 2.55 for the hybrid algorithm.
differences in performance of load-balanced versions of CMT-nek on Vulcan. The original code did not finish in 5 hours.

Figure 7.4. Performance comparison between load-balanced and original versions of CMT-nek on Vulcan. They were run on 65,536 MPI ranks, that is 16,384 nodes with 4 cores per node. Adaptive distributed load balancing was used. The average time per time step for the original and the load balanced versions was 20.00 and 2.52 seconds, respectively, giving us an overall speed-up of 7.9. Original code did not finish in 5 hours.

Figure 7.5. Performance comparison between adaptive load balanced and user-triggered load balanced versions of CMT-nek on Vulcan. They were run on 32,768 MPI ranks, which is 8,192 nodes with 4 cores per node. Distributed load balancing algorithm was used. The time per time step for the user-triggered and the adaptive load-balanced versions was 4.17 and 3.78 seconds, respectively, for the last 2,000 steps, giving us an overall improvement of 9.4%.
version didn’t finish in 5 hours. The time per time step for the original and the load-balanced versions was 20.00 and 2.52 seconds, respectively, giving us an overall speed-up of 7.9. Load balancing happened before simulation time step 1. There was no need to load balance after that since the time per time step didn’t increase over the threshold set to trigger load balancing.

Figure 7.5 shows a comparison between the adaptive load-balancing and user-triggered load-balancing algorithms. For the user-triggered load-balancing algorithm, the \( k = 500 \), thus load balance is triggered every 500 time steps. As we can see from the figure, there is no performance degradation in the first 4,000 time steps, making any load balancing redundant during this time. However, right after step 4,000 performance degrades sharply, requiring frequent load balancing. The user-triggered load-balancing algorithm is insensitive to these performance variations and continues to load balance every 500 time steps. The average time per time step from step 4,000 to step 6,000 taken by the adaptive and user-triggered load balancing versions was 3.78 and 4.17 seconds, respectively. Thus, adaptive load-balancing algorithm gained an overall improvement of 9.4% compared to the user specified triggered load-balancing algorithm, and further the load balancing happens automatically without requiring any intervention by the user.

7.3 DVFS on CMT-nek

Energy efficiency and power minimization have become critical for high performance computing systems. Most modern processors and co-processors are equipped with dynamic voltage and frequency scaling (DVFS) mechanisms, where the operating frequency of a processor can be changed to lower power consumption. Additionally, only a subset of processors can be used to save overall energy when the scaling of the application does match with the increase in power requirements. In this work, we develop a systematic approach for deriving energy-performance trade-offs on a hybrid multi-core (CPU+GPU) processor. Using CMT-bone, we show how to derive energy-performance tradeoffs on a server consisting of multi-core and multiple GPU processors based on studies involving individual types of cores. The search space consisting of a combination of number of CPU cores, number of GPUs, CPU frequency, GPU frequency and loading factor on CPUs and GPUs, make an exhaustive experimental study impossible. Instead we experimentally determine the Pareto optimal configurations for each type of cores and combine them to estimate the Pareto optimal configurations for the whole system. Figure 7.6 shows the estimated and actual energy for points on a Pareto optimal curve obtained using our dynamic programming method. The configurations are given in the order of CPU load, CPU count, GPU count, CPU Frequency, GPU Frequency, and Energy. The Figure shows how close our prediction is to the actual results. In our paper, we further detail the methodology and error analysis we have done as part of this project.
7.4 CMT-bone on KNL

7.4.1 KNL Architecture

We start by describing our Intel KNL System. We used Intel Xeon Phi Processor 7210 for our experiments. It has 64 cores where each core has four hyper-threads. The operating processor base frequency is 1.30 GHz and L2 cache size is 32 MB. The Thermal Design Power (TDP) is 215 Watts which means that when all cores are operating at base frequency under a high complexity workload that fully utilizes the cores then the average power consumption will be 215 Watts for the chip.

7.4.2 Hybrid Computing Performance (MPI + OpenMP)

We found using prior experiments, that using 64 MPI processes on KNL gives us the best performance. The total number of spectral elements for the test we tried was 8192 and the total number of particles was 512,000. Auto-vectorization was enabled in all the runs to benefit from AVX512 instructions. Next, we added OpenMP support to CMT-bone to be able to utilize hyperthreading. We mainly parallelized the do loops using OpenMP parallel loop directives. This reduces the communication overhead between processes and improves performance due to shared memory. Thus, we started with 64 MPI processes and used hyperthreading to enable the OpenMP
threads. As shown in Figure 7.7, we obtained a speedup of 10% when 64 MPI processes with 2 OpenMP threads per process. The Y-axis shows the running time for one simulation time step.

![Figure 7.7. Comparison of using different number of OpenMP threads on CMT-bone on KNL.](image)

7.4.3 Performance using Different Memory Modes

To analyze the effect of MCDRAM on the performance of CMT-bone, we collected performance results without using MCDRAM, with MCDRAM in cache mode and with MCDRAM in flat mode. The performance speedup from MCDRAM usage reaches 1.73X compared to not using MCDRAM when all the 64 cores of Intel KNL are being utilized.

7.4.4 Performance using different Clustering Modes

KNL has three basic clustering modes. All-to-all, quadrant and hemisphere, and SNC4 and SNC2. Among these All-to-all is the default. Figure 7.8 shows the effect of using different MCDRAM and clustering modes on CMT-bone performance over Intel KNL.

As we can see from Figure 7.8, using MCDRAM in Flat mode yields the best performance. This is because our application fits completely in the 16 GB MCDRAM even for a large input size. For example, we have 8,192 spectral elements in total which is 128 elements per process when all the 64 cores are used. This is a load on the larger side for the CMT-bone application. For other applications, if the processes consume more that 16GB memory then other configurations such as cache only mode may give better performance.
7.4.5 Comparison among Different Architectures

The architectures used in this comparison are shown in Figure 7.9. This includes two architectures with GPUs, an Intel Broadwell and a KNL 7210. We found that on Intel Broadwell, the best performance was obtained using 32 MPI processes. Hence, we used 32 MPI processes on Intel Broadwell, and 64 MPI processes on Intel KNL. On GPU architectures, only one MPI process is used to run on the entire GPU. Figure 7.10 shows the performance comparison between different platforms for the optimal number of MPI processes for Intel Broadwell and Intel KNL.
architectures. Thus, we see that Intel KNL gives us the best performance. The speedup for KNL is 1.5x compared to Intel Broadwell and 1.2x compared to Tesla P100.

Figure 7.10. Performance comparison of CMT-bone on different platforms.
8. Exascale Team

8.1 Overview

The Exascale research focuses on the development and application of Behavioral Emulation (BE), a coarse-grained simulation method, to enable rapid design-space exploration (DSE) of design strategies, optimizations, and parallelizations on extreme-scale systems up to Exascale. Such exploration supports the CCMT center by providing insight into how to optimize CMT-nek for potential candidate architectures before those architectures exist.

In this annual report, we will focus on the achievements in the following areas:

1. BE methods and tools
2. “Closing the loop” - application of BE methods and tools for CMT-nek DSE
3. FPGA acceleration of Behavioral Emulation

8.2 Exascale Behavioral Emulation

The Exascale Behavioral Emulation (BE) research team focuses on the use of BE methods to enable rapid design-space exploration (DSE) of design strategies, optimizations, and parallelization on extreme-scale systems up to Exascale. In Behavioral Emulation, low-level details are abstracted away to improve scalability, while retaining sufficient accuracy for design-space exploration. Such exploration supports the project by providing insight into how to optimize CMT-nek for potential existing, near-future, or notional architectures. A concept diagram illustrating the BE co-design process is shown in Figure 8.1.

Figure 8.1. Co-design using Behavioral Emulation.

In this section of the annual report, we will report our progress for this past year and the plans on the following tasks:
Center for Compressible Multiphase Turbulence

- BE-SST simulator and its use for CMT-bone-BE simulations
- Multi-fidelity surrogate (MFS) model for performance prediction
- BE simulation of CMT-nek design space
- Lab setup for power/energy/thermal modeling
- FPGA-acceleration methods for rapid DSE space reduction & UQ

We first give a brief update on the development of the BE-SST simulator and tools, followed by the use of BE-SST for coarse-grained simulation of CMT-bone-BE on existing architectures for validation; and for performance prediction on notional architectures. We will describe how BE is combined with MFS for low-cost model validation. The use of BE simulation is illustrated with a case study of design space exploration of CMT-nek. Finally, our initial work of incorporating power/energy/thermal modeling into BE is described, followed by our progress on using FPGAs to accelerate behavioral emulation for rapid DSE.

8.2.1 BE – SST and its use for CMT-bone-BE simulations

After establishing BE-SST simulator as our primary simulation platform, the main objective this past year was to achieve better scalability while also improving accuracy and performance. In this reporting period, we redesigned the existing interpolation models; and also communication models, transitioning it from static to dynamic routing, thus improving the scalability of the simulator. BE-SST, with these extensions, was then used to perform validation of our simulations on existing architecture and performance predictions (in terms of execution time) on notional architectures.

First, the performance models of BE-SST were updated to support the performance models generated by symbolic regression, in addition to interpolation techniques (which was used previously).

For communication models, non-blocking communication operations such as ISEND and IRECV were also implemented. Then, a transition from static routing to dynamic routing was carried out. The dynamic routing eliminated the construction of a routing table for each message. The performance of the simulator with dynamic routing was studied on systems up to one million cores. The system's build time, which was earlier a bottleneck (with static routing) for large simulations, was reduced with dynamic routing. Figure 8.2 depicts this performance scale-up of the simulator which was tested on a system ranging from 64 cores to one million cores, connected in a 3-D mesh topology running CMT-bone-BE. These simulations were performed on 64 ranks (2 nodes) on HiPerGator, a high performance computing system at the University of Florida. Since there is no separate build time for the routing information, the system's build time was not a hindrance in the performance of the simulator.
With the need to scale-up to Exascale architectures, the necessity to simulate bigger systems is indispensable. However, scaling up led to increased memory consumption. With increase in simulation size, the memory distribution became unbalanced between the MPI Ranks. Rank 0 consumed the largest proportion of Graph Construction Memory (GCM). We are working with Jeremey Wilke from Sandia National Labs to fix this uneven memory distribution.

The capability of simulator to simulate CMT-bone-BE up to a million-core system was demonstrated and is shown in Figure 8.3. We simulated scaled-up versions of Vulcan and Titan (let us call Vulcan+ and Titan+, respectively) as a first step towards notional architectures using existing machines. This gives us an idea of how an application would have performed if there was a system bigger than Vulcan or Titan. The simulations were validated up to 128k cores on Vulcan and Titan, shown as red circles in Figure 8.3. Prediction results from the simulator are indicated by blue circles. The average percentage error between CMT-bone-BE and BE simulation of Vulcan and Titan is 4%, thus proving good modeling accuracy.

We wanted to simulate larger than 1M core system but due to the memory distribution issues described earlier, we had to restrict the simulation to 800k cores. As we are stepping into notional architectures, we have planned to simulate the performance of the new ORNL system,
SUMMIT’18. As SUMMIT uses non-blocking fat-tree topology, we have segregated the implementation into 3 steps: 1) implementation of fat-tree architecture in BE-SST; 2) implementation of a dynamic routing mechanism; 3) making it non-blocking. Currently, support for fat-tree architecture has been added to the simulator. With limited information on the exact specifications of SUMMIT’18, support to several types of fat-tree architectures (dual-rail fat-tree, single-rail fat-tree, folded-clos fat-tree) is being implemented.

Going forward for BE-SST, we will implement dynamic routing scheme for the fat-tree and converting the fat-tree into non-blocking fat-tree. We will also enhance the communication modules of the simulator. The enhanced BE-SST will be used to explore different notional architectures to illustrate the capability of the simulator. Documentation of the simulator’s code is being done and will be completed in the next quarter.

8.2.2 Multi-fidelity surrogate models for performance prediction

Previously, we collaborated with the Uncertainty Budget (UB) team to reduce the discrepancy between CMT-bone and BE simulation data by fitting BE simulation to CMT-bone using Multi-Fidelity Surrogate (MFS). We also worked on interpolation schemes for BE-SST and studied noise in HiPerGator, a high performance computing system at UF.

This past year, we used MFS with BE to perform low-cost model validation of the parent application, CMT-nek. Parent applications are computationally very expensive to run, and thus slows down the process of design space exploration (DSE). Hence, for rapid DSE, mini-apps are used. To use these mini-apps for co-design, the mini-app needs to be validated against its parent app, which is an expensive process. To reduce this cost, we proposed a multi-fidelity surrogate (MFS) approach which combines data samples of low-fidelity (LF) models (representative apps and architecture simulation) with a few samples of a high-fidelity (HF) model (parent app).

For the case study shown in Figure 8.4, we had 22 CMT-nek data points and 125 BE simulation data points in the design space. 10 runs (out of 22) were selected randomly and fixed as the validation runs and the others were used as training data. With less than 10 data samples of CMT-nek (high fidelity HF data) and all the 125 data points of BE simulation (low fidelity LF data), the MFS significantly reduces the error and provides a good fit to CMT-nek. The RMSE error between predictions of MFS and the baseline HF models was 4% alone as shown in Figure 8.4, which is significantly better than using either LF or HF data. This enables us to perform validation of the parent app at reduced computational budget. A paper was published to the PMBS’17 workshop at the Supercomputing Conference 2017, where MFS is proposed as a method to reduce computational budget of parent application.
We extended this work to incorporate extrapolation into the MFS framework. Extrapolation is the methodology in which we use both high fidelity (HF) and low fidelity (LF) data points in the design space to predict a single data point or a collection of data points farther away from the design space. In our case, we predicted CMT-nek up to 1 million processors as shown in Figure 8.5. BE was used to predict CMT-bone-BE execution time on Vulcan up to a million processors, which was used as LF data, and CMT-nek runs up to 128k processors where used as HF data. The concept was first validated using CMT-nek data within the design space. Upon validation, extrapolation proved accurate with less than 10% error in prediction. Figure 8.5 also shows the prediction of CMT-nek execution time for a million processors along with 95% confidence interval (C.I.).

![Figure 8.4: RMSE error (left) and R-MD error (right) between CMT-nek validation runs and MFS predictions based on BE simulation.](image)

Figure 8.5. Predicting CMT-nek execution time using MFS over entire design space (left) and for one million processors with 95% C.I. (right).
Going forward, we will extend the MFS paper into a journal paper by including extrapolation of large scale runs to it. We are also planning to study co-kriging framework for MFS and include the results in the journal paper.

8.2.3 BE simulation of CMT-nek design space:

A primary goal of Behavioral Emulation (BE) is to help develop CMT-nek for the future systems, which we achieve via algorithmic design-space exploration (DSE) of CMT-nek on notional Exascale architectures.

In the past year we have seen significant changes in the CMT-nek code, particularly in the particle-solver phase where lots of particle-tracking capabilities were added to CMT-nek. To perform DSE in the algorithm front, we need to identify the key kernels in CMT-nek which have a major role in the application performance. CMT-nek contains two main phases: Gas solver and Particle-solver. We began our DSE efforts by profiling the application to identify the potential candidates which would benefit most from algorithmic DSE. Based on our interaction with the code-development team, we determined the following parameter space for profiling as shown in Table 8.1.

<table>
<thead>
<tr>
<th>Application parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element-size (lx1)</td>
<td>5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25</td>
</tr>
<tr>
<td>Elements/core (lelt)</td>
<td>8, 16, 32, 64, 128, 256</td>
</tr>
<tr>
<td>Particles/gridpoint (α)</td>
<td>0.1, 0.33, 1, 3.33, 10</td>
</tr>
</tbody>
</table>

Figure 8.6 shows the profiling results for various parameter combinations. Initially, for small input sizes, the Gas solver is the dominant kernel in CMT-nek application. However, as the input size increases, the Particle solver scales non-linearly and quickly becomes dominant. As a result, we started looking into alternate algorithms available in the Particle solver kernel. The Particle-solver calculates the particle properties during the course of the simulation. It does so by calculating the forces acting upon each particle by the surrounding gases at each timestep and updates the particle position accordingly. The current implementation uses one-way coupling to calculate the particle forces. It uses an rk3 (3 stage Range-Kutta) time-integration algorithm and barycentric interpolation to calculate the fluid/gas properties acting upon each particle. The code development team suggested few alternatives for time-integration and interpolation stages, i.e., bdf (Backward differentiation formula) and reduced-barycentric interpolation. Both these algorithms are less expensive than their counterparts but lose some accuracy due to simplification. In this study, we are interested the performance vs. accuracy trade-off for these set of algorithms.
Figure 8.6. Execution time of Gas solver and Particle solver kernels with varying element-size and particles/gridpoint (α).

Table 8.2 and Figure 8.7 show the parameter space and results of BE-simulations of the Particle solver for rk3 and bdf algorithms, respectively. Note that the execution-time is only the computation time as there is no communication done for the given problem specification. Also,
we generated the BE models for various kernels in the particle-solver phase using multi-variate linear regression for interpolation for multiple-parameters (> 2), which is expensive and inefficient.

Table 8.2: Parameter set for BE-Simulation and Validation for CMT-nek particle solver.

<table>
<thead>
<tr>
<th>Application parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element Resolution (lx1)</td>
<td>5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25</td>
</tr>
<tr>
<td>Elements-per-processor (lelt)</td>
<td>4, 8, 16, 32, 64, 128, 256, 512</td>
</tr>
<tr>
<td>Particles/gridpoint (α)</td>
<td>0, 0.1, 0.33, 1, 3.33, 10</td>
</tr>
</tbody>
</table>

Figure 8.7. Comparison of Particle solver execution time per timestep (computation) using rk3 and bdf time integration algorithms by (a) varying element size(lx1) (b) varying particles/gridpoint(α) and (c) varying element count (lelt).
Figure 8.7 shows the BE-simulations results of the particle-solver kernel for different time-integration algorithms. Based on the simulations, we observed that both time-integration algorithms vary non-linearly w.r.t. element size (lx1) and elements-per-processor (lelt) and vary linearly w.r.t. particles/gridpoint (α). The bdf algorithm provides 3x speedup compared to Range-Kutta. Similarly, reduced-barycentric interpolation provides an average speedup of 5x compared to barycentric interpolation and reaches a theoretical maximum of 8x for larger element sizes.

Apart from the algorithmic design space, we also need to consider the parameter design space. CMT-nek is a complex code and with the inclusion of the Particle solver, it only becomes more complex. Particle data consumes a significant portion of the processor memory and with the increase of number of particles per core, it soon runs out of memory. For example, for an element size (lx1) of 15, and with 15 elements-per-processor (lelt), if there are no particles (i.e., α = 0) CMT-nek consumes a memory of 0.6GB. If we include 1 particles/gridpoint (α = 1), memory consumption goes close to 1GB and with 3 particles/gridpoint (α = 3), CMT-nek consumes a total memory of 1.5 GB. The number of particles present in a core/problem is an application-specific parameter which may vary drastically based on the type of simulation problem we’re solving. Therefore, we need a memory consumption model which provides total memory consumed by CMT-nek based on the key input parameters, i.e., element size, elements-per-processor and particles/gridpoint. Such a model also helps in determining the maximum problem size we can run on a HPC machine before encountering a memory shortage issue.

Table 8.3: Parameter space used for building CMT-nek memory consumption model

<table>
<thead>
<tr>
<th>Application parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element Resolution (lx1)</td>
<td>5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19</td>
</tr>
<tr>
<td>Elements-per-processor (lelt)</td>
<td>4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096</td>
</tr>
<tr>
<td>Particles/gridpoint (α)</td>
<td>0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
</tbody>
</table>

To generate memory-consumption model, we started by collecting memory-consumption data for varying values of the input parameters. The parameter space is shown in Table 8.3. CMT-nek is written in Fortran 77, which does not allow dynamic memory allocation. Hence, the total memory consumed is obtained by considering the size of executable generated after compilation. Using our in-house benchmarking automation framework, we generated scripts which compile CMT-nek for various combinations of lx1, lelt and α, as shown in Table 8.3. We collected approximately 800 data points and performed a multi-variate linear regression to generate a memory-consumption model. We have eliminated the insignificant order trailing terms from the model. The CMT-nek memory consumption model (in bytes) is shown as follows:

\[
CMT - \text{nek}_{\text{mem}} = [2960 + 1065\alpha](lx1)^3(lelt) + [7372 + 5537\alpha](lx1)^2(lelt) + [4044 + 200\alpha](lx1)^3
\]
As mentioned earlier, we used multi-variate linear regression to generate performance models for BE simulations and also for generating memory-consumption model. However, there are certain limitations to multi-variate linear regression. The main limitation is that in order to generate accurate performance models, we need to have proper knowledge of the kernel we are modelling. In other words, we need to specify the base equation of how the model performs and linear regression helps in determining the coefficients. Such knowledge is not readily available at all times. We were considering other alternate approaches for building BE models during the latter half of the year and eventually decided to use symbolic regression. To perform symbolic regression, we leveraged a genetic-programming based symbolic regression tool that we developed in another research project. This symbolic-regression tool takes as input definitions of the function parameters, along with training data, and then performs a lengthy exploration to discover an equation that minimizes error in the training set. Table 8.4 shows a comparison of error metrics between symbolic regression and linear regression, with symbolic regression reducing mean absolute error by 126.37% on average.

<table>
<thead>
<tr>
<th>Particle solver kernel</th>
<th>Linear Regression</th>
<th>Symbolic Regression</th>
<th>Error Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAPE</td>
<td>RMSE</td>
</tr>
<tr>
<td>Compute1</td>
<td>4.4281E-6</td>
<td>3.5040%</td>
<td>1.08E-5</td>
</tr>
<tr>
<td>update_particle_location</td>
<td>1.2580E-5</td>
<td>10.82%</td>
<td>1.26E-5</td>
</tr>
<tr>
<td>particles_in_nid</td>
<td>1.17E-2</td>
<td>76.35%</td>
<td>1.27E-2</td>
</tr>
<tr>
<td>interp_props_part_location</td>
<td>4.511E-1</td>
<td>902.60%</td>
<td>3.672E-1</td>
</tr>
<tr>
<td>upd_vel_and_posrk3_stage1</td>
<td>8.2186E-4</td>
<td>399.0318%</td>
<td>5.43E-4</td>
</tr>
<tr>
<td>upd_vel_and_posrk3_stage3</td>
<td>1.0648E-4</td>
<td>19.3695%</td>
<td>6.5E-5</td>
</tr>
<tr>
<td>upd_vel_and_posrk3_allstage</td>
<td>7.3443E-3</td>
<td>27.3041%</td>
<td>7.57E-4</td>
</tr>
<tr>
<td>upd_vel_and_pos_bdf</td>
<td>5.3822E-4</td>
<td>20.03%</td>
<td>1.21E-3</td>
</tr>
<tr>
<td>red_interp</td>
<td>9.33E-2</td>
<td>4.6895%</td>
<td>2.9445E-2</td>
</tr>
<tr>
<td>usr_particles_forces</td>
<td>8.7998E-5</td>
<td>6.2690%</td>
<td>6.85E-5</td>
</tr>
<tr>
<td>tri_interp</td>
<td>1.5843E-4</td>
<td>1.3893%</td>
<td>5.34E-5</td>
</tr>
</tbody>
</table>

Table 8.4: Comparison of error metrics between linear and symbolic regression of particle-solver kernels.

Also, in the last quarter, we began profiling the particle-solver communication. As previously mentioned, the particle-solver kernel calculates the forces acting upon the particle and updates the particle properties at each time-step. These properties include position, velocity etc. However, due
to the particle's movement, a particle may move from one processor domain to another processor. In order to facilitate particle communication, CMT-nek uses the *crystal-router* algorithm, which completes the particle movement among N processors in log(N) stages. However, the particle movement is not static, as it depends on problem-specific parameters; hence the amount of data being transferred between processors is dynamic and changes with each time-step. To simulate such a data-dependent kernel, we perform trace-driven simulations: we collect the trace of particle locations at every time-step and feed it to a tool we developed to generate useful information required by the BE-SST simulator. We call this tool the Particle-Distribution tool. The tool outputs two key results: one is the number of particles belonging to each processor at a given time-step (which specifies the workload on each processor), and the other is the number of particles moving from a processor's domain to its neighbors (specifies the communication workload). As the particle movement is problem-specific and doesn’t depend on number of processors, we need only a single trace to observe the particle-distribution on even multiple processors.

Going forward, we will work on updating the symbolic regression tool so that it provides a stochastic distribution instead of a deterministic value for the simulated time. Also, we will generate the BE models of the Gas solver phase in order to perform a full-scale BE simulation of the CMT-nek application.

8.2.4 Lab setup for power/energy/thermal modeling

With power consumption and thermal issues being a major constraint on Exascale computing, we have begun the process of integrating power, energy, and thermal modeling into Behavior Emulation, as illustrated by the green circles in Figure 8.8. With these models, we will be able to explore Pareto-optimal tradeoffs between performance, power consumption, energy, and thermal issues during BE simulations. We will then be able to use this information to perform design-space exploration to optimize CMT-nek for different potential Exascale architectures.

![Figure 8.8. Energy modeling for Behavioral Emulation.](image-url)
For this energy-modeling research, we had previously used IR thermal imaging of a single core AMD desktop computer processor as a proof-of-concept for identifying hot spots on a processor. This processor was stripped down to just its silicon dye and was then placed back on its motherboard to be recorded by our FLIR camera as it booted up. We could visibly see the hot spot on the processor where the core was located to show that hot spots will pop up when a core is being used more than other parts of the processor. From this we have moved onto using an in-house Intel Knights Landing processor that consists of 64 cores. This server uses fans on the side to blow air across instead of the typical overhead fan which will allow us to take thermal images from above without altering the server layout much. However, we are still hindered by the metal heat sink on this server as it is currently too dangerous to remove the heat sink on such an expensive processor, as opposed to the old desktop processors we were using that could be easily be stripped down to its lowest layers.

To be able to get images of the processor while running, we needed to cut a hole in the casing. The system will not power on without being in the casing and the casing blocks our visual access to the processor. With the help of the Mechanical Engineering Department at UF, we cut a hole right above the processor that granted us visible access to the processor while it runs and allow us record thermal data. Upon cutting this hole the first step was to start running benchmarks, collecting data, and seeing what we had to work with to help us gain an idea of how we want to move forward. We ran CMT-nek on the processor in 5-minute periods while using a different number of cores during each period to be able to observe differences in heat patterns and variations in the temperature. We could see what we were expecting in that we were using much more heat across the entire processor while running all 64 cores while we also saw the heat was somewhat grouped up in certain areas of the processor when only running 8, 16 and 32 cores. Some of the photos taken with the IR camera are shown in Figure 8.9.

![Figure 8.9: Thermal images showing delta T of the KNL processor running CMT-nek. One core is used in the left image while 64 cores are used in the right image.](image-url)
Moving forward, we want to learn more about what we can do to reduce the maximum temperatures we measure while the processor is running while continuing to improve our experimental set up. To do so, we will start by using OpenMPI to gain an understanding of how the KNL is doing its mapping by running examples to draw out this information. Once we understand how this mapping is done by the server, we can then essentially start to “inject” heat patterns across cores by selecting certain cores to run computationally intensive code, while the others sit idle. Essentially, we will be using OpenMPI machine files to bind processes to specific cores and observe the thermal patterns we are able to read. We will verify these heat patterns with the FLIR camera to make sure we are seeing heat where we would expect it to. From this we can create metrics and start experimenting with load balancing to reduce maximum temperatures and overall energy consumption.

8.2.5 FPGA-acceleration methods for rapid DSE space reduction & UQ

The BE approach is based on distributed discrete-event simulation (DES) between BE objects (BEOs). BEOs mimic the abstract behavior of applications or architectures using high-level surrogate models to predict each event’s impact on the targeted simulation objectives (e.g., performance, reliability, power). Although this approach enables faster simulation than traditional cycle-accurate approaches, Exascale simulation with existing tools could take minutes, hours, or...
even days to complete a single simulation. These lengthy simulations place very practical limits on DSE and Uncertainty Quantification (UQ) efforts that often require thousands, or even millions of independent simulations. To address this issue, we have been exploring an FPGA-accelerated approach based on pipelined DES that focuses not necessarily on improved performance for a single simulation, but instead on increased simulation throughput. By focusing on throughput, we unlock the potential for huge performance gains when the problem under study calls for numerous independent simulations (e.g., DSE, Monte Carlo simulation).

In our approach, we designed a custom compiler to convert the MPI parallel application and architecture specification used as input to existing simulation tools (Figure 8.10(a)) to a data-flow graph (DFG) representation (Figure 8.10(b)). Graph vertices represent each unique discrete event (e.g., matrix multiply, MPI send, barrier) and edges represent their input/output dependencies. The compiler utilizes several graph-optimization techniques to manipulate the DFG before ultimately mapping it to an FPGA pipeline (Figure 8.10(c)). Assuming sufficient resources, the compiler simply maps each vertex operation of the DFG to independent FPGA resources. By adding pipeline registers between events with dependencies in the DFG, a single simulation has a latency equivalent to the DFG’s critical path. More importantly, successive simulations start/complete once every cycle. However, if required resources exceed a single FPGA (a near certainty for Exascale simulation) some degree of resource sharing and/or circuit partitioning across multiple FPGAs is required. There exist many approaches to this problem (the focus of current and future research) and the challenge becomes how to apply them to automatically generate Pareto-optimal circuits for any simulation. We have developed a resource-sharing strategy that attempts to “collapse” the essentially 2D DFG (threads by events per thread) into a 1D pipe with two major scaling advantages: (1) resources scale sublinearly with the number of threads (with a factor of “threads” cost in simulation throughput), and (2) pipelines scale as a single, unidirectional pipe that can be partitioned predictably across any number of connected FPGAs with only minimal overhead.

Table 8.5 presents performance data for the two approaches while highlighting the advantages and disadvantages of each. The no-sharing approach clearly has superior performance in terms of simulation throughput and latency, but consumes far more resources and is ultimately far less scalable. Comparing lines 1-9 with lines 10-16 & 17-23, resources scale linearly with timesteps (TS) for both approaches, but the sharing approach allows for many more TS due to its much lower base utilization. Additionally, the sharing approach's resource utilization scales sublinearly with the number of threads (1, 10, & 17) allowing for simulations far larger than previously possible with the no-sharing approach on a single FPGA (e.g., simulated configurations up to 2,147,483,648 ranks). Although the sharing approach appears to suffer a significant drop in performance when considering simulation throughput (inverse proportionality to ranks), if we instead consider event throughput we see direct proportionality to logic utilization (LU%) almost independent of the number of ranks. Simulation throughput decreases as the amount of work per simulation increases (increased events per simulation as number of ranks increase); but the amount of work completed each clock cycle remains constant and depends upon how much event hardware is instantiated.
This ultimately means that although simulation throughput will continue to decrease with increased ranks, event throughput will remain relatively constant dependent on LU.

Overall, the proposed approach provides simulation/event throughput that is many orders-of-magnitude faster than the BE-SST software simulator (speedup column shows at least 6 orders), however, this gain in throughput comes at a cost. One notable limitation of the FPGA-pipelined approach is a sacrifice in analysis capabilities largely due to limited I/O bandwidth (e.g., BE-SST can log all intermediate event data for postmortem analysis while the pipelined approach is limited to a handful of “monitored” events that can be logged without causing pipeline stalls). We also note that there are several potential application behaviors that the FPGA may not be able to efficiently handle at the same level of granularity as software (e.g., dynamically modifying control flow based on event timing). One possible solution is to model the application at a higher level of abstraction such that the behaviors are no longer present. For our envisioned use case, these limitations are not prohibitive because a designer can use our approach to rapidly prune a huge design space into a small set of promising candidates that can then be explored in more depth using existing techniques.

We have tested collapsed pipeline configurations up to 2,147,483,648 threads and continue to observe the same scalability trends. Going forward, we are currently working on a CMT-bone.BE
example that spans across a configurable number of FPGAs in Novo-G#. With this example, we plan to test the scalability of this approach for more than a single FPGA. We are also working on several other example applications (e.g., LULESH, select HPC Dwarfs) codes to test this approach with. In the process of developing these codes we have encountered a few application behaviors that we cannot currently express with the existing BE tools, which has prompted us to add this functionality. We are currently integrating support for performance models generated by symbolic regression (described in Section C). This involves creation of a tool that auto generates performance model pipelines based on the provided symbolic regression equations and modification of the existing BE-FPGA compiler to integrate these autogenerated pipelines into the larger BE simulation pipeline. We plan to showcase this functionality with an end-to-end case study where we consider a huge design space, probably containing hundreds of thousands of parameter and algorithmic combinations and prune the design space to a few tens or hundreds of possible optimal candidates with FPGA simulation. The remaining candidates are then analyzed more closely with BE-SST.
9. Deep Dives

9.1 Exascale Deep-dive

The University of Florida held a Deep Dive Workshop on Feb 3-4, 2015. The Agenda is presented below, and the presentation slides can be found on the CCMT webpage at:

https://www.eng.ufl.edu/ccmt/events/workshops/

Agenda:

Deep Dive
University of Florida
February 3-4, 2015

Current Attendee List:

Bob Voigt                  NNSA HQ                  rvoigt@krellinst.org
Matt Bement                LANL                  bement@lanl.gov
David Daniel              LANL                  ddd@lanl.gov
Dave Nystrom               LANL                  wdn@lanl.gov
Maya Gokhale               LLNL                  maya@llnl.gov
Martin Schulz             LLNL                  schulzm@llnl.gov
Jim Ang                   SNL                  jaang@sandia.gov
Arun Rodrigues             SNL                  afrodr@sandia.gov
Jeremy Wilke            SNL                  jjwilke@sandia.gov

S. Balachandar “Bala”       University of Florida  bala1s@ufl.edu
Alan George                University of Florida  george@hcs.ufl.edu
Rafi Haftka                University of Florida  haftka@ufl.edu
Herman Lam                University of Florida  hlam@ufl.edu
Sanjay Ranka              University of Florida  ranka@cise.ufl.edu
Greg Stitt                  University of Florida  gstitt@ece.ufl.edu
Tom Jackson                University of Florida  tlj@ufl.edu

Tania Banerjee          University of Florida  tmishra@cise.ufl.edu

University of Florida Students:
Dylan Rudolph               rudolph@hcs.ufl.edu
Nalini Kumar                nkumar@hcs.ufl.edu
Carlo Pascoe                carlo.pascoeg@gmail.com
Kasim AlliKasim            kasimalli490@yahoo.com
Chris Hajas                  chrisnijas@ufl.edu
Mohammed Gadou             mgadou@ufl.edu
Michael Retherford

95
UF Deep dive agenda:

**Tuesday, February 3, 2015**

8:20        Van pickup at Hilton
8:30 – 9:00  Breakfast
9:00 – 9:30  Welcome and Deep-Dive Overview (3 Sessions)
   1. Behavioral emulation (BE): modeling & simulation/emulation methods
   2. CS issues (performance, energy, and thermal)
   3. Use of reconfigurable computing to accelerate behavioral emulation

* Each of the three deep-dive sessions is designed to be interactive: a combination of short presentations by UF and Tri-lab researchers, intermixed with discussion, demonstrations, etc.

9:30 – 11:30  Session 1: Behavioral Emulation: Modeling & Simulation/Emulation Methods
   • UF topics:
     o Behavioral characterization
     o Parameter estimation
   • Tri-lab topics:
     o Overview of FastForward 2 and DesignForward 2 (Jim Ang, SNL)
     o Multi-scale architectural simulation with the Structural Simulation Toolkit (Arun Rodrigues, SNL)

11:30 – 12:30 Lunch
12:30 – 2:00  Session 1 (continued): Behavioral Emulation: Beyond Device Level
   • UF topics:
     o Synchronization for speed
     o Congestion modeling
     o Behavioral characterization & modeling beyond device level
   • Tri-lab topics:
     o Using discrete event simulation for programming model exploration at extreme-scale (Jeremy Wilke, SNL)
     o ASC next-generation code projects (David Daniel, LANL)

2:00 – 5:00  Session 2: CS Issues (Performance, Energy, and Thermal)
   • UF topics:
     o Performance and autotuning for hybrid architectures
     o Energy and thermal optimization
     o Dynamic load balancing
   • Tri-lab topics:
     o Performance, energy, and thermal benchmarking (Jim Ang, SNL)
     o Why power is a performance issue: utilizing overprovisioned systems (Martin Schulz, LLNL)

* There will be an afternoon coffee break in this time slot

6:30        Dinner (University Hilton)

**Wednesday February 4, 2015**
9.2 Multiphase Physics Deep-dive

The University of Florida held a Multiphase Physics Deep Dive Workshop on October 6-7, 2016. The Agenda is presented below, and the presentation slides can be found on the CCMT webpage at:

https://www.eng.ufl.edu/ccmt/events/workshops/

Agenda:

**Multiphase Physics Deep-dive Workshop**

**October 6-7, 2016 Attendee List**

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georges Akki</td>
<td><a href="mailto:gakiki@ufl.edu">gakiki@ufl.edu</a></td>
<td>University of Florida</td>
</tr>
<tr>
<td>Subramanian Annamalai</td>
<td><a href="mailto:subbu.ase@ufl.edu">subbu.ase@ufl.edu</a></td>
<td>University of Florida</td>
</tr>
<tr>
<td>Marco Arienti</td>
<td><a href="mailto:marient@sandia.gov">marient@sandia.gov</a></td>
<td>Sandia</td>
</tr>
<tr>
<td>S. Balachandar</td>
<td><a href="mailto:balal1s@ufl.edu">balal1s@ufl.edu</a></td>
<td>University of Florida</td>
</tr>
<tr>
<td>Ankur Bordoloi</td>
<td><a href="mailto:ankur@lanl.gov">ankur@lanl.gov</a></td>
<td>LANL</td>
</tr>
<tr>
<td>Alexander Brown</td>
<td><a href="mailto:albrown@sandia.gov">albrown@sandia.gov</a></td>
<td>Sandia</td>
</tr>
<tr>
<td>Jesse Capecelatro</td>
<td><a href="mailto:jcaps@umich.edu">jcaps@umich.edu</a></td>
<td>University of Michigan</td>
</tr>
<tr>
<td>Seungwhan Chung</td>
<td><a href="mailto:schung58@illinois.edu">schung58@illinois.edu</a></td>
<td>University of Illinois</td>
</tr>
<tr>
<td>Paul Crittenden</td>
<td><a href="mailto:pcritte@ufl.edu">pcritte@ufl.edu</a></td>
<td>University of Florida</td>
</tr>
<tr>
<td>William Dai</td>
<td><a href="mailto:dai@lanl.gov">dai@lanl.gov</a></td>
<td>LANL</td>
</tr>
<tr>
<td>Angela Diggs</td>
<td><a href="mailto:angela.diggs.1@us.af.mil">angela.diggs.1@us.af.mil</a></td>
<td>Eglin AFB</td>
</tr>
<tr>
<td>Timothy Dunn</td>
<td><a href="mailto:dunn13@llnl.gov">dunn13@llnl.gov</a></td>
<td>LLNL</td>
</tr>
<tr>
<td>Brad Durant</td>
<td><a href="mailto:neoncrash@ufl.edu">neoncrash@ufl.edu</a></td>
<td>University of Florida</td>
</tr>
<tr>
<td>John Eaton</td>
<td><a href="mailto:eatonj@stanford.edu">eatonj@stanford.edu</a></td>
<td>Stanford University</td>
</tr>
<tr>
<td>Mahdi Esmaily</td>
<td><a href="mailto:mesmaily@stanford.edu">mesmaily@stanford.edu</a></td>
<td>Stanford University</td>
</tr>
</tbody>
</table>
Center for Compressible Multiphase Turbulence

Giselle Fernandez  
gisellefernandez@ufl.edu  
University of Florida

Marianne Francois  
mmfran@lanl.gov  
LANL

Joshua Garno  
jgarno@ufl.edu  
University of Florida

Jason Hackl  
jason.hackl@ufl.edu  
University of Florida

Alan Harrison  
alanh@lanl.gov  
LANL

Jeremy Horwitz  
horwitz1@stanford.edu  
Stanford University

Kyle Hughes  
kylehughes89@ufl.edu  
University of Florida

Thomas Jackson  
jlj@ufl.edu  
University of Florida

Rahul Koneru  
rahul.koneru@ufl.edu  
University of Florida

Allan Kuhl  
kuhl2@llnl.gov  
LLNL

Ali Mani  
aliman@stanford.edu  
University of Florida

Yash Mehta  
ymehta@ufl.edu  
University of Florida

Chandler Moore  
wcm0015@tigermail.auburn.edu  
University of Florida

Brandon Morgan  
morgan65@llnl.gov  
LLNL

Balu Nadiga  
balu@lanl.gov  
LANL

Fady Najjar  
najjar2@llnl.gov  
LLNL

Brandon Osborne  
bosborne3@ufl.edu  
University of Florida

Fred Ouellet  
f.ouellet@ufl.edu  
University of Florida

John Parra-Alvarez  
jcpp@lanl.gov  
University of Utah

Katherine Prestridge  
bertrandrollin@gmail.com  
Embry-Riddle

Kevin Ruggirello  
krarelling@lanl.gov  
LLNL

Kambiz Salari  
salar1@llnl.gov  
Sandia

Shane Schumacher  
scsschum@sandia.gov  
Sandia

Philip Smith  
philip.smith@utah.edu  
University of Utah

Sean Smith  
sean.t.smith@utah.edu  
University of Utah

Prashanth Sridharan  
shani130@ufl.edu  
University of Florida

Jeff St. Clair  
jeff.stclair@navy.mil  
University of Florida

Cameron Stewart  
csstewart10@ufl.edu  
University of Florida

Jeremy Thornock  
jthornock@gmail.com  
University of Utah

Gretar Tryggvason  
gtryggva@nd.edu  
University of Notre Dame

Markus Uhlmann  
markus.uhlmann@kit.edu  
Karlsruhe Institute of Tech.

Laura Villafane  
lvillafa@stanford.edu  
Stanford University

Robert Voigt  
rvoigt@krellinst.org  
ASC AST

Seng Keat Yeoh  
syeoh2@illinois.edu  
XPACC University of Illinois

Duan Zhang  
dzhang@lanl.gov  
LANL

Ju Zhang  
jzhang@fit.edu  
Florida Institute of Tech.
Agenda Multiphase Physics Deep-Dive, October 6-7, 2016
St. Petersburg Marriott Clearwater,
12600 Roosevelt Blvd, North St. Petersburg, FL 33716, Phone: 727-572-7800

Thursday Oct 6th, 2016
7:30 – 8:30 Breakfast (Provided)
8:30 – 8:45 Welcome (S. Balachandar)
8:45 – 10:15 Overviews
  Gretar Tryggvason Challenges and opportunities in fully resolved simulations of multi fluid flows
  Ali Mani Overview of computational modeling at Stanford PSAAP: particle-laden flows subject to radiative heating
  Kambiz Salari Research activities for energetic dispersal of particles
10:15 – 10:30 Coffee
10:30 – 12:00 Overviews
  Alex Brown Multiphase Methods for Modeling Fire Environments
  Phil Smith Multi-phase flow modeling at Utah PSAAP – predictivity in application
  S. Balachandar Overview of multiphase flow computational strategy at UF PSAAP
12:00 – 12:15 Further Discussion
12:15 – 1:30 Lunch (Provided)
1:30 – 3:00 Macroscale
  Duan Zhang Equations and Closures for Deformation and Flow of Continuous and Disperse Materials
  Allen Khul 3-Phase Model of Explosion Fields
  Marco Arienti Multiphase Flow Simulation Strategies at the CRF
3:00 – 3:15 Coffee
3:15 – 5:15 Microscale
  Jeremy Horwitz Point-particle modeling for two-way-coupled problems: Challenges, verification, and physics-based improvements
  Georges Akiki Extended point particle model
  Jesse Capecelatro Recent insights on turbulence modeling of strongly-coupled particle-laden flows
  Tom Jackson Microscale simulations of shock particle interaction
6:30 – 9:00 Dinner (Provided – all attendees)

Friday Oct 7th, 2016
7:30 – 8:30 Breakfast (Provided)
8:30 – 10:30 Modeling & numerical methods
   - Sean Smith: Particle dynamics: coal-specific modeling
   - A. Subramaniam: Microscale modeling based on Generalized Faxen theorem
   - Alan Harrison: Modeling of Ejecta Particles in the FLAG Continuum Mechanics Code
   - Markus Uhlmann: Large scale microscale simulations and modeling opportunities
10:30 – 10:45 Coffee
10:45 – 12:15 Experiments and simulations
   - Ankur Bordoloi: Experimental measurements of drag on shocked particles
   - Laura Villafañe: Including real experimental effects in validation of numerical models for confined particle-laden flows
   - Fady Najjar: Meso-scale Simulations of Shock-Particle Interactions
12:15 – 1:30 Lunch (Provided)
1:30 – 3:00 Meso/macroscale
   - Mahdi Esmaily: A systematic study of turbophoresis by four-way-coupled simulation of Stokesian particles in channel flow
   - Balu Nadiga: Bayesian Analysis of Inter-Phase Momentum Transfer in the Dispersed Eulerian Formulation of Multiphase Flow
   - John Parra-Álvarez: Eulerian Models and Polydispersity Treatment for Dilute Gas-Particle Flows
3:00-4:00 Discussion and Closing remarks
10. Publications

2014


2015


2016


2017


2017

2018
11. Conferences and Presentations

2014


2015


2016


2017


Center for Compressible Multiphase Turbulence


12. Workshops Held or Attended

2014
2. Dr. Tania Banerjee. Salishan Conference, April 2014.
5. Heather Zunino presented tomographic PIV at a PIV workshop to a group of faculty, post-docs, and Ph.D. students at Instituto Superior Técnico (IST) in Lisbon, Portugal. She was invited to stay for this workshop and present after meeting with several researchers from IST at the Laser Symposium this summer.

2015
10. Heather Zunino presented tomographic PIV at a PIV workshop to a group of faculty, post-docs, and Ph.D. students at Instituto Superior Técnico (IST) in Lisbon, Portugal. She was invited to stay for this workshop and present after meeting with several researchers from IST at the Laser Symposium summer 2015.

2016
12. Dr. Sanjay Kumar and Ph.D. students Kasim Alli and Carlo Pascoe (2016). Attended the WEST workshop. Dr. Kumar gave a talk.
15. Rahul Koneru and Fred Ouellet attended “Scaling your science on Mira” workshop, May 24-25, Argonne National Laboratory.

2017

13. Students and Staff Internships
13.1 Internships Completed

2014
1. Heather Zunino, Ph.D. Student (ASU), US, Dr. R. Adrian. Ms. Zunino completed her 10-consecutive week stay at Los Alamos National Laboratory, under the mentorship of Dr. Kathy Prestridge this summer (May-August 2014). Dr. Prestridge is the Extreme Fluids Team leader in the Physics-23 group at LANSCE. Project: Vertical Shock Tube (Calibration and Tomographic PIV), Horizontal Shock Tube (Particle Tracking Program).
2. Kevin Cheng, MS Student, Florida. Lawrence Livermore National Laboratory. Mentor: Dr. Maya Gokhale, Dr. Scott Lloyd. Project: An Emulation Framework for Tracing near Memory Computation. US, Dr. Alan George, ECE, MS (graduated Fall 2014), core.

2015

5. Dr. Bertrand Rollin, Staff Scientist, CCMT. March 16-20, 2015, LANL.

6. Nalini Kumar, Ph.D. Student, India, ECE, Dr. Alan George, part cost share and part leveraged. (Internship not required). March-May, 2015. LLNL. Dr. James Ang.

7. Chris Hajas, M.S. Student, US, ECE, Dr. Herman Lam, core. May 18-August 18, 2015 at LLNL with Dr. Maya Gokhale.

8. Christopher Neal, Ph.D. Student, US, MAE, Dr. S. Balachandar, core. June 14-August 20, 2015 at LLNL with Dr. Kambiz Salari. Chris’ work was highlighted in the Stewardship Science Academic Programs (SSAP) Annual, which highlights the exceptional research activities and people supported by the SSAP (PSAAP is under the SSAP).

9. Carlo Pascoe, Ph.D. Student, US, ECE, Dr. Herman Lam, core. Will intern summer 2015 at LLNL with Dr. Maya Gokhale.


11. Dr. Tania Banerjee, PhD., one week internship in May, 2015 to LLNL.

12. Justin Thomas Mathew, MS Student (2015), Drs. Haftka and Kim, core. Visited Los Alamos National Laboratory. T6-Theoretical Biology and Biophysics group, Dr. Nick Hengartner, studying and developing extensions of epidemiological mathematical models of infectious disease.

2016

13. Dr. Chanyoung Park, Postdoc, CCMT. Feb 22-26, 2016 at LLNL with Dr. Samuel P. Schofield.

14. Dr. Jason Hackl (Postdoc, CCMT) visited Sandia National Labs, Albuquerque, New Mexico from January 16-20 to present CMT-nek to Greg Weirs and other personnel.


17. Georges Akiki, Ph.D. Student, US, MAE, Dr. Balachandar, core. Internship to LANL. May-August, 2016, Dr. Marianne Francois.

18. Dr. Tania Banerjee, PhD., one week internship in June, 2016 to Sandia.

2017

19. Maria Giselle Fernandez-Godino, PhD Candidate, CCMT. May 06- July 26, 2017 at University of Toulouse, Toulouse France with Dr. Christian Gogu.


22. Mohamed Gadou, Ph.D. Candidate, CISE, Dr. Ranka, core. Internship to LANL Summer, Dr. Galen Shipman.
24. Yash Mehta, Ph.D. Student, MAE, Dr. Balachandar. Internship to LLNL Summer, 2017, Dr. Kambiz Salari.

13.2 Internships Planned
1. Prashanth Sridharan, Ph.D., US, MAE, Dr. Balachandar, core, Internship to LANL Summer 2018.
2. Fred Ouellet, Ph.D. Student, US, MAE, Dr. S. Balachandar, core

13.3 Internships Not Yet Planned
3. Yiming Zhang, Ph.D., China, MAE, Drs. Haftka and Kim, cost share (internship not required)
4. Rahul Koneru, Ph.D., India, MAE, Dr. Balachandar, cost share (internship not required)
5. Joshua Garno, Ph.D. Student, US, MAE, Dr. Balachandar, core
6. Sam Nili, Ph.D. Student, US, MAE, Drs. Haftka and Kim, core
7. Brad Durant, Ph.D., MAE, Dr. Balachandar, core
8. Ryan Blanchard, BS, US, ECE, Dr. Herman Lam, core
9. Keke Zhai, Ph. D. student, China, Dr. Ranka, CISE, UF, cost share (internship not required)
10. Aravind Neelakantan, ECE, Dr. Herman Lam, cost share, (internship not required)
11. Sai Prabhakar Rao Chenna, ECE, Dr. Herman Lam, cost share, (internship not required)

13.4 Graduated Students
1. Kevin Cheng, MS Student (2014). Dr. Alan George, ECE.
4. Angela Diggs, Ph.D. (2015). US, MAE, Dr. S. Balachandar (other funding, internship not required). Currently employed at Eglin AFB and working with the Center.
8. Parth Shah, M.S. Student (2016), India, ECE, Dr. Herman Lam.
11. Sankeerth Mogili, M.S. (2017). Dr. Ranka, CISE, UF
13. Nalini Kumar, Ph.D. (August, 2017), Dr. H. Lam, ECE. Intel, Santa Clara CA

13.5 Placement of Staff
1. Dr. Bertrand Rollin, Staff Scientist, CCMT. Faculty position at Embry-Riddle, Fall 2015.
2. Dr. Mrugesh Shringarpure, Postdoc, CCMT. Researcher, Exxon Mobil, Spring 2016.
4. Dr. Georges Akiki, PhD (2016), Dr. S. Balachandar, MAE; Postdoc thru March 2017, Currently Postdoctoral Associate, LANL.

14. NNSA Laboratory Interactions

2014
1. Rob Cunningham, LANL. Setting up "Florida" file sharing group on Mustang
2. Blaise Barney, LLNL. Setting up account on Vulcan and Surface.
4. Nathan Fabian, SNL. Introduction to Catalyst.
5. Don Frederick, LLNL. Issue with submitting a run on Vulcan.
7. Jan Nunes, LLNL. Account request on Edge
8. Discussions with Donald Frederick of Lawrence Livermore National Laboratory related to MPI issues on the Vulcan computing cluster – discussion was relevant to parallel operation of the Rocflu code.
9. Telecon with Paraview Catalyst developers and users (Greg Weirs, Nathan Fabian, Kenneth Moreland at Sandia National Laboratory) at about deploying the Catalyst library into Rocflu for in-situ visualization.
10. Worked with Greg Lee (Livermore Computing Center)–who is a software debugging/troubleshooting expert at LLNL–to get Rocflu to run at scale on the Vulcan computing cluster at LLNL.
11. Interactions with Rich Cook (Livermore Computing) who is the visualization expert at LLNL–we are working with him to get in-situ visualization capabilities using Catalyst integrated into Rocflu.
12. Interacted with David DeBonis at Sandia National Laboratories to get setup with using PowerInsight for power and energy measurements.
13. Interacted with Robert Cunningham and Amanda Bonnie at Los Alamos National Laboratory for temperature measurements using the data collection tool Lightweight Distributed Metric Service.
14. Interacted with Justin A. Too and Daniel J. Quinlan at the Lawrence Livermore National Laboratory on compilation and installation of ROSE.
15. Interacted with Blaise Barney at the Lawrence Livermore National Laboratory on various system issues, including access to CVS.
16. Road trip to SNL (Albuquerque) and LANL, March 24 – 26, 2014 (Herman, Bala, and Rafi)
17. SNL (Albuquerque: Jim Ang, Arun Rodrigues, Scott Hemmert, Simon Hammond - from SST team at SNL, Albuquerque, NM
18. SNL (Livermore): Jeremiah Wilke from SST Macro team at SNL, Livermore, CA

2015
19. LLNL: Maya Gokhale regarding proposal to analyze memory access demands of CMT-Nek kernels and evaluate the potential benefits of utilizing stacked memories with embedded logic functions to increase performance.

20. Dr. Steve Beresh (SNL), visit CCMT and gave talk, Thursday April 23, 2015.

21. The members of the microscale group (in particular Chris Neal and Yash Mehta, another graduate student) have worked closely with the staff at LLNL's Livermore Computing center to resolve an issue that Rocflu was having when it was scaled up and executed on the BG/Q machine Vulcan. We have also been in touch with Rich Cook at LLNL's Livermore Computing to continue our work with Paraview's Catalyst library integration into Rocflu. During the fall and spring, Chris Neal has been in touch with Dr. Kambiz Salari to keep him up-to-date on the progress of Chris's research. Chris and Yash were aided by Blaise Barney (LLNL) in renewing their computing accounts at LLNL.

22. Drs. Jason Hackl and Mrugesh Shringarpure and David Zwick and Goran Marjanovic. Visited Argonne National Lab for the 2015 Nek User/Developer meeting to commit CMT-nek to the nek5000 repository and work intensively with Prof. Paul Fischer.

23. Barry Rountree, LLNL. Tania worked closely with Barry’s group to come up to speed on using RAPL to measure and bound power on Intel platform.

24. Tanzima Islam, LLNL. Tania is working with Tanzima on validating CMT-bone proxy app.

25. David DeBonis, SNL. Tania interacted with David to come up to speed on using PowerInsight for power and energy measurements.

26. Patrick Lavin, a summer intern with Barry Rountree, worked on the performance issue and used vectorization effectively on the derivative optimization code getting about 40% improvement using vectorization alone. Tania will follow up with verifying if it is possible to get further improvement with CHiLL based autotuning approach.

27. S. Balachandar and H. Lam attended and presented at the 2015 ASC PI meeting at Monterey, CA.


2016

29. S. Balachandar and S. Ranka attended and presented at the 2016 ASC PI meeting at Las Vegas, NV.

30. Tania Banerjee collaborated with Jeremy Wilke, Gwen Voskuilen and Arun Rodriguez on MLM.

31. Tania Banerjee collaborated with Tanzima Islam to validate proxy app CMT-bone.

32. Justin Mattews and Chanyoung Park collaborated with Justin Wagner (Sandia).

33. Kyle Hughes collaborated with Kathy Prestridge (LANL) in sharing p-rad experiments of rapidly dispersed pre-fragmented particles.

34. Jeff St. Clair and Balachandar are interacting with Fady Najjar (LLNL) in performing ALE3D simulations of intense shock propagation over a close-packed bed of deformable aluminum particles.

35. Balachandar interacted with John Schwarzkf (LANL) in completing a review article on compressible multiphase flow for the second evolution of multiphase flow handbook.

36. Balachandar and Jackson organized and hosted the Deepdive workshop on multiphase flows. Approximately 15 research scientists from the three NNSA laboratories attended.
37. Yash Metha interacted with Kambiz Salari (LLNL) on three-dimensional simulations of a strong shock propagating through a random bed of particles. This work has been written up and is in the final stages of submission as an archival journal paper.
38. Tom Jackson, Fady Najjar (LLNL), and Habib Najm (Sandia) are organizing a Focus Topic session on UQ in high speed flows for the upcoming APS SCCM meeting in July.
39. The UQ team has regular teleconference calls with Greg Weirs (Sandia).

2017
41. T.L. Jackson, visited with Dr. Mark Short, LANL, August 2017.
42. The UQ team has regular teleconference calls with Greg Weirs (Sandia).
43. S. Balachandar and T.L. Jackson interacting with Dr. F.M. Najjar on fully resolved inviscid simulations of flow over a particle close to a wall.
44. S. Balachandar attended the 2017 ASCI PI meeting and presented at Monterey, CA
45. Jeff St Clair and Balachandar are using the Navy code Dismas along with the Sandia solid mechanics code Peridynamics to solve the flow structure interaction problem at detonation conditions.
46. Yash Meta is continuing his interaction with Dr. Kambiz Salari, in both processing his shock propagation over a random array of particles and process the data from shock propagation over a moving array of particles.
47. From Sandia Kevin Ruggirello and Shane Schumacher attended David Zwick’s PhD proposal in September 2017 via teleconference.
48. Dr. Jason Hackl visited Sandia National Laboratory January 16-19, 2017, and presented. His interactions with Drs. Bill Rider, Travis Fisher, Greg Weirs and John Shadid were a wealth of education on testing, assessing convergence, and crucial issues in entropy stability of high-order schemes.
49. Dr. Jason Hackl attended the 5th International Workshop on High-Order CFD Methods, AIAA SciTech, Kissimmee, FL, January 6-7, 2018, and met with Dr. Travis Fisher (SNL) to discuss artificial viscosity and entropy splitting.
50. Chanyoung Park: Met Vicente Romero from Sandia and discussed the idea of cost based adoptive sampling and presented the Sandia shocktube validation and UQ to him at the SciTech 2018 conference.
51. Kyle Hughes: Met Justin Wagner (Sandia) and Daniel Guildenbecher (Sandia) and shared his experience on the P-rad experiment and the Eglin tests at the SciTech 2018 conference.
52. Justin Mathew: Met Justin Wagner (Sandia) and received feedback about our UQ study for the Sadia shock-tube experiments at the SciTech 2018 conference.
53. Herman Lam and students meet with Maya Gokhale at SuperComputing 2017.