Cover Page

Proposal

FOA Number: DE-FOA-0002501
US Department of Energy Office of Science Advanced Scientific Computing Research
DOE/Office of Science Program Office Technical Contact: Dr. William Spotz, ASCR
William.Spotz@science.doe.gov

PAMS Preproposal number: PRE-0000027087

Multiscale Data Reduction for Analysis and Visualization of Extreme-Scale Particle Data via Scientific Machine Learning

Lead Principal Investigator

Han-Wei Shen, Professor
The Ohio State University
Computer Science and Engineering
395 Dreese Lab, 2015 Neil Aenue, Columbus, OH 43210
Phone 740-417-6645, Email shen.94@osu.edu

Co-Principal Investigators

Ayan Biswas, Computer Scientist, Los Alamos National Laboratory
Christopher Fryer, Computer Scientist, Los Alamos National Laboratory
Jonah Miller, Computer Scientist, Los Alamos National Laboratory

Submitted by

The Ohio State University
395 Dreese Lab 2015 Neil Avenue
Columbus, OH 43210
Administrative Contact: Patricia Evans, Grant Administrator
Phone: 614-292-3149, Email: evans.377@osu.edu

Contents

1 Introduction

As scientists eagerly anticipate the benefits of exascale computing, our limited ability to explore the vast amount of data at scale has become a major roadblock to further accelerating scientific discovery. The efficiency of data movement does not scale with the rapidly increasing volume of data, thus the relative cost of data movement will continue to increase. Computational scientists must therefore transform their simulation outputs to reduced representations with salient features preserved to obtain insight and perform accurate spatiotemporal analysis. In this project, the collaborative team will pioneer a comprehensive strategy that focuses on hierarchical spatiotemporal analysis of particle data to perform data transformation, reduction, and super-resolution reconstruction. Many DOE applications such as astrophysics, cosmology, molecular dynamics, and combustion use mesh-free approaches that require particles as their basic primitives. This presents a greater challenge in data reduction and feature extraction, as compression algorithms such as SZ and ZFP find it more difficult to compress point clouds compared to regular-grid structured data. In this project, we will develop novel machine learning techniques to perform data transformation to represent particle data by a latent representation (or particle latent in short), which is a succinct but information-rich representation, based on their spatial and physical attributes. From the transformed representations, we will perform data reduction to represent particle data hierarchically at different scales, which can result in data of several orders of magnitude smaller in size. To allow scientists to perform detailed analysis of the particle data at highest possible quality, we will develop techniques to generate spatiotemporal super-resolution of the reduced representation to reconstruct data at significantly better quality than existing lossy compression algorithms. Finally, to quantify the amount of information content represented by the particle latent and measure the reconstruction error, we will develop uncertainty quantification methods based on the information theory. This will enable an understanding of the relationship between the input, latent, and the reconstructed output. Our approaches will be applicable to a wide range of applications in computational sciences that produce particle data.

1.1 Potential Scientific Impact

The proposed research will have a direct impact on many scientific applications that routinely generate very large scale, time-varying particle data. Currently, a number of particle-based methods are being actively developed for a broad set of applications. To test the efficacy of our methods, we will apply them to three applications: Smoothed particle hydrodynamics (SPH), Monte Carlo radiation transport (MC), and tracer particles. All three domains benefit significantly from high-fidelity, multi-resolution, compression of particle data.

SPH uses a set of Lagrangian ”interpolation points” to model hydrodynamics, retaining many of the strengths of Lagrangian methods without the issues of mesh entanglement [52]. A “smoothing” kernel is used to interpolate between SPH particles so that hydrodynamic quantities can be calculated across the entire space in the simulation. Because of its ability to trace mass (and hence cover a wide range of length scales) and its natural ability to conserve mass and momentum, SPH has been applied to a wide number of engineering and astrophysical applications[63, 62, 55, 53, 38, 70, 7, 52]. Current computational power can facilitate SPH calculations in excess of 100 million particles.

MC radiation transport models the six-dimensional Boltzmann equation by evolving discrete sample particles, each of which represents a radiation packet, such as a collection of photons or neutrinos. Here we focus on neutrino transport in compact object mergers, where it has recently enabled breakthroughs in our understanding of these systems [61, 20, 49, 50, 51, 21]. A major advantage of MC methods over approximate transport methods is that the full six-dimensional distribution function is available. Modern transport applications routinely use tens to hundreds of millions of MC particles.

Tracer particles are a Lagrangian diagnostic tool often utilized in Eulerian methods. In an Eulerian method, the field of interest—usually a fluid—is evolved on a grid. Tracer particles are simulated particles that are carried by the fluid flow. Each tracer provides a history for a Lagrangian fluid packet, evolved by the grid Eulerian code. (Some recent applications can be found in [50, 51].) Although the number of tracer particles is relatively modest compared to other applications—of order 1 million particles—tracer data is often output at high cadence, meaning tracer data can consume hundreds of gigabytes for a single simulation.

1.2 Research Overview

The goal of this project is to develop scientific machine learning based approaches for an end-to-end particle data reduction and reconstruction pipeline, where the components of the pipeline can not only seamlessly work with each other, but also stand on their own to serve specific purposes. As shown in Figure 1, our research contains three major components: (1) particle data transformation, (2) multiscale analysis, reduction, and super-resolution, and (3) uncertainty-aware particle reconstruction. The goal of particle data transformation is to distill important information from particles and organize the particles in a way to make it suitable for adaptive coarsening and refinement. Using distribution-based data aggregation and advanced deep learning models, the multiscale analysis, reduction, and super-resolution component is designed to perform hierarchical data assimilation, aggregation, reduction, and ultimately high quality data reconstruction. Finally, the uncertainty-aware particle reconstruction component allows scientists to reconstruct particles of superior quality from the reduced data and stay informed with the accuracy and uncertainty of the reconstructed data. Because deep learning models are heavily used in this research, an information-theoretic approach will be taken to analyze the quality of the models and quantify the amount of information that flows through the deep learning model layers. We note that while this proposal is not focused on in situ data analysis, the transformation of particle data and hierarchical data analysis and reduction can be performed at simulation time (in situ) once the machine learning models are trained. Then spatiotemporal super-resolution and uncertainty-aware data reconstruction are performed in the post-hoc analysis stage. Below, we provide a brief overview of our research methods.

I. Particle Data Transformation: With the advent of neural networks, methods such as autoencoders have been developed to extract compact latent representations which can encode the most salient features of the data. Creating particle latent information by encoding the distributions of both spatial and physical attributes around each particle's local neighborhood allows scientists to focus on the essential geometric and physical structure of their data, promising more effective feature extraction and tracking. Although there exist many latent generation methods, extracting latent representations from meshless particle data is more challenging. In this research, we will develop fundamentally new methods based on neural networks to extract latent representations for particles that are several orders of magnitude smaller in size while retaining all the salient information in the data sets.

Particle data that are meshless present greater challenges for data reduction and feature extraction due to the lack of effective methods for interpolation and filtering. It is also computationally more expensive to retrieve particles that are in close proximity and therefore extracting coherent structures at different scales is non-trivial. To enable multiscale data analysis and facilitate hierarchical data reduction, we propose methods to transform the positional information of particles to discrete displacement fields which will lend themselves well to a wider range of methods for data aggregation, adaptive coarsening, and spatial and temporal super-resolution. The discrete displacement fields are analyzed and aggregated jointly with the particle latent representations in such a way that multiscale analysis and reconstruction can be performed.

II. Multiscale Analysis, Reduction, and Super-Resolution: The transformed particle data in the forms of particle latents and discrete displacement fields pave the way for effective multiscale data analysis and reduction. Research will be conducted to develop adaptive data reduction schemes based on user selected quality requirements to aggregate the particle displacements and latents at different scales. Distribution-based methods will be developed to represent the reduced data such that it is possible to preserve salient data features and in the mean time present the error and uncertainty to scientists. Since it is imperative to

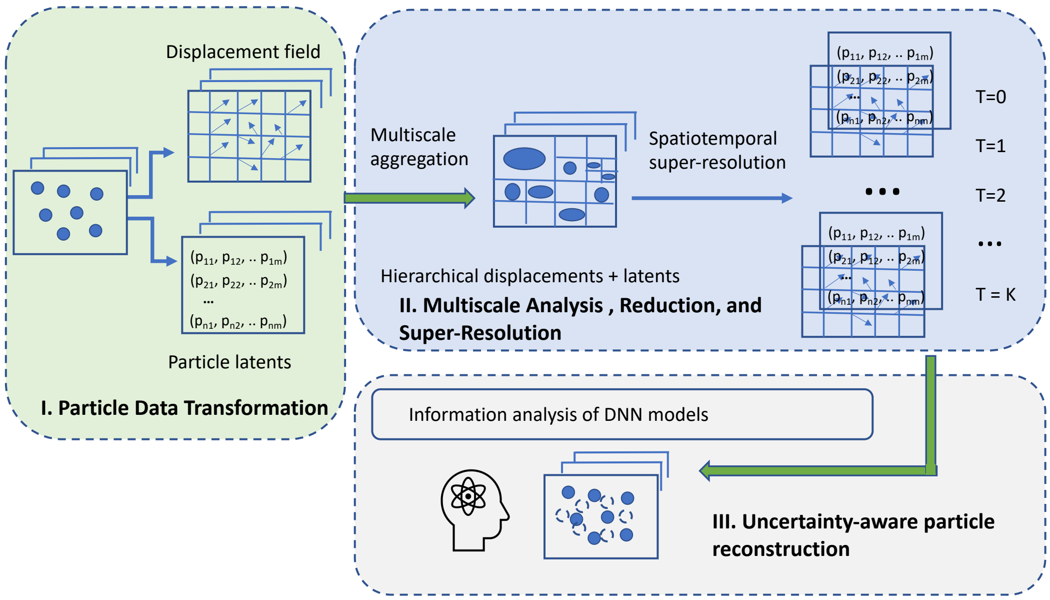

Table 1: Overview of research methods.

provide scientists with data of the best quality for post-hoc analysis and visualization, fundamental research in spatial and temporal super-resolution methods that can reconstruct particle data of superior quality from the reduced representations will be conducted. Although several machine learning based super-resolution techniques have been developed, none of them are designed for particle data, and most of them focus on performing scaling of the entire domain at a fixed rate, with no concern in the existence of features and the control of trade-off between quality and storage consumption. In this research, we will develop deep neural networks (DNNs) to perform spatiotemporal super-resolution reconstruction from reduced particle data. To adapt to the Lagragian nature of particles, we will train neural networks to perform motion estimation and temporal fusion to recover the missing information in time. We will also leverage multi-scale Generative Adversarial Networks (GANs) to allow arbitrary scaling factors in different spatial subdomains.

III. Uncertainty-aware Particle Reconstruction: To allow the scientists to fully control the quality of data for post-hoc analysis, it is important to provide them with a clear indication of errors when data are reduced and reconstructed. In our particle data reduction pipeline, there are two sources of error. One is the error associated with multiscale data aggregation, and the other is the error related to data reconstruction via super-resolution deep neural networks. Since we will employ distribution-based modeling to aggregate the multi-dimensional particle attributes, we will develop methods tailored to the distribution models that will be in use. For the spatiotemporal super-resolution components, error estimation and analysis will be performed for the deep neural network models. Furthermore, we will use information-theoretical approaches to analyze the information flow and the loss of quality through different layers of the neural network models. Finally, to fully quantify the uncertainty of data reconstruction, we will use a Bayesian approach to quantify the likelihood of having particles reconstructed at particular positions with specific attribute values.

1.3 State of the Art

1.3.1 Particle Visualization and Point Cloud Neural Networks

Particle simulations are used across a plethora of scientific fields. Analyzing and understanding this kind of data is a hot topic, as evidenced by the fact that 2015, 2016 and 2019 IEEE Scientific Visualization Contests all focused on particle datasets [14, 24, 67]. For these datasets, the tasks of feature extraction and tracking are critical for understanding the development of the simulated phenomena over time. Recent work [60] also proposed a method to interactively visualize large scale particle datasets by representing them with Gaussian

Mixture Models (GMMs). Meanwhile, researchers in the computer vision domain have designed and applied neural networks [58, 57, 80, 30] to point cloud data and have achieved great success for classification and segmentation; both require meaningful feature extraction from point clouds.

A deep learning survey for 3D point clouds provides a comprehensive overview of the recent point cloud neural network studies [26]. The PointNet [57] is one of the first methods in deep learning for point cloud data. Subsequently, Qi et al. expended the PointNet with PointNet++ by applying the PointNet recursively on a nested partitioning of the input point cloud to improve efficiency and robustness [58]. PointNet++ borrows the ideas of hierarchically applying filters to extract both local and global features from Convolutional Neural Networks(CNN). Lots of recent works improve on the basic architecture of PointNet++ [30, 80, 73, 40, 85, 42, 32]. GeoConv [40] is another neural network that shows good performance with relative small network size. LassoNet [12], which solves the problem of interactive selection of objects with lasso is one of the few visualization works we can find that utilizes 3D point cloud neural networks.

1.3.2 Super-Resolution for Scientific Data

There are two categories of super-resolution: spatial super-resolution (SSR) and temporal super-resolution (TSR). SSR aims to increase the spatial resolution of input data. Zhou et al. [89] use a 3D convolutional neural network (CNN) to perform SSR on volumes with better feature reconstruction than trilinear interpolation or cubic-spline interpolation. Guo et al. [25] use three neural networks in parallel to perform SSR on 3D vector fields. Xie et al. create tempoGAN [87], which upscales fluid flows for temporally consistent high resolution output. Fukami et al. compare two ML-based SSR methods for 2D fluid flow [22]. SSR-TVD by Han and Wang [27] uses a generative adversarial network (GAN) for upscaling volumes such that the upscaled output frame is temporally coherent with adjacent timesteps. A similar line of research to SSR in the computer vision community is called image super-resolution (ISR), and we refer readers to a survey by Wang et al. [81] for recent techniques. With the recent advances in super-resolution, Jakob et al. have created a public 2D flow dataset with varying Reynold's number, citing that ML methods are a powerful method for interpolating flow maps and data reduction [35].

TSR aims to increase the temporal resolution of input data. TSR-TVD by Han et al. [28] uses a convolutional LSTM to learn the recurrence between timesteps efficiently to recover interpolated timesteps with higher accuracy than linear interpolation. Other methods perform both spatial and temporal super-resolution together with spatiotemporal super resolution (STSR). Fukami et al. [23] performs STSR on fluid flow data using two networks in a 3D+1D approach, where the spatial dimension is upscaled first, and then the temporal dimension. MeshfreeFlowNet [36], proposed by Jiang et al., trains a network to allow inference for arbitrarily sized spatiotemporal domains with PDEs impacting training through a loss function.

1.3.3 Neural Networks for Visualization

Most existing research on AI for scientific visualization focuses on single-run data. More than a decade ago, Ma et al. [46, 76, 77] pioneered the use of traditional neural networks (nowadays called "shallow," as opposed to modern "deep" neural networks) for classifying multivariate volume data for 3D visualization. With the explosive growth of modern deep learning techniques, the community seeks to use the capabilities of neural networks to address various scientific visualization problems. Recently, researchers have studied the use of DNNs for volume upscaling [90], transfer function refinement for visual feature matching [59], complex volumetric structure depiction and exploration [13], creation of a generative model for volume rendering [8], high-resolution and perceptually authentic image synthesis [34], viewpoint quality estimation [68], and isosurface upscaling in image space [84]. Our team developed InSituNet [31], an image-based surrogate model to support parameter space exploration by synthesizing visualization results for given simulation parameters and won the best paper award for IEEE Visualization 2019.

2 Research Methods

2.1 Transformation of Particle Data

A particle data set typically consists of the positional information (x, y, z), an identifier, and multiple physical attributes for each particle. This raw format of data is often challenging for data compression or feature extraction as the order used to store the particles is not always correlated to the structure of coherent regions or features. Because of the lack of meshes that connect the particles, it is more computationally expensive to identify the neighbors of each particle, which plays a crucial role in removing redundancy while preserving salient features. To tackle this challenge, there is a need to perform data transformation for particle data in such a way that the properties of particles in local neighborhoods can be easily extracted and exploited for data reduction and feature preservation and extraction. In the following, we describe our approaches for transforming particle data.

2.1.1 Latent Representations for Particle Attributes

A particle represents a discrete sample of the underlying continuous field, but it is the coherent structures sharing some common physical properties in the domain that make the candidates of salient features and allow for effective data reduction. In this research, our goal is to produce a succinct but information-rich representation, called latent representation, for particle data by taking each particle's own attributes and those of the particles in its neighborhood into account. The goal of the latent representation is to compactly capture the high-level structure of the particle field, such as the distance between particles and the distributions of physical attributes in the neighborhood of the particles. To do this, we propose to use a deep learning approach that employs particle autoencoders to create a latent representation for each particle. To train the autoencoder, unlike image data, particle data do not have meshes to connect the points, hence the commonly used convolutional neural networks (CNN) for autoencoders cannot be easily applied. In this research, the candidates for us to investigate include PointNet [58, 57], GeoConv [40], and LassoNet [12]. PointNet and its related extensions are a generalization of the convolutional neural network to particle data, in which the convolution is applied to a group of particle positions, and pooling produces the latent representation; GeoConv [40] tackles similar types of problems but shows good performance with relative small network size. Finally, LassoNet [12], whose goal is to solve the problem of interactive selection of objects with lasso and is one of the few visualization works that utilize 3D point cloud neural networks.

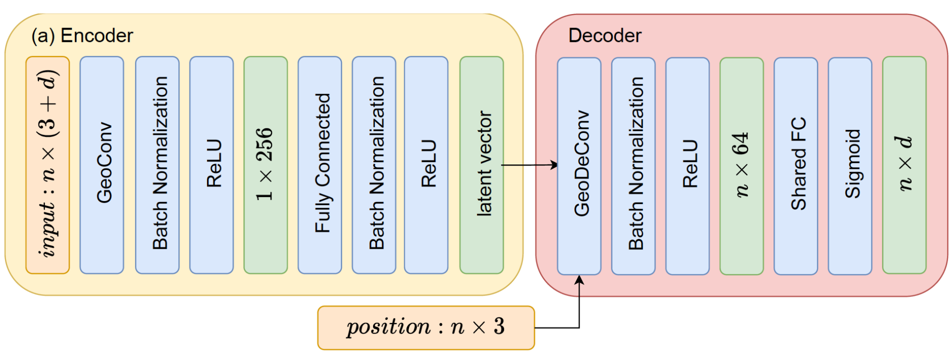

Figure 1: The architecture of our autoencoder. The input is given in a particle patch with n particles, each of which has 3 dimensional position and d dimensional attributes.

Recently we have performed preliminary sttudies using GeoConv[40] to extract latent representations from particle data sets[41]. Given a particle p, the core idea is to project the coordinates of neighboring particles q into the orthogonal directions of particle p and define a kernel on those directions. We aggregate

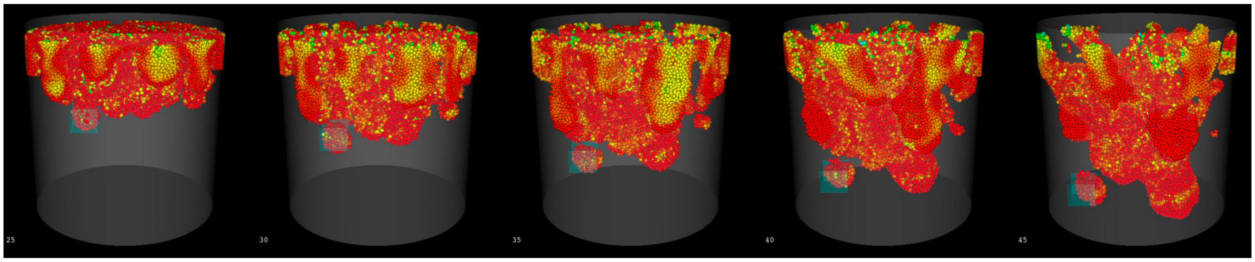

Figure 2: Examples of particle reconstructed from their latent representations in a time sequence.

the attributes of particle q in p's neighborhood into p based on its energy in different directions as:

$$g(\vec{p},\vec{q}) = \sum_{\vec{b}\in B} \cos^2(\vec{pq},\vec{b}) W_{\vec{b}} X_{\vec{q}} \tag{1}$$

, where $\vec{b}$ is one of the orthogonal bases vectors, $W_{\vec{b}}$ is the learnable weights in the corresponding direction, and $X_{\vec{q}}$ is the attributes for particle q. We can calculate a vector representation for particle p by aggregating all particles in the neighborhood according to their distance to the center:

$$\text{GeoConv}(\vec{p},r) = \frac{\sum_{\vec{q}\in N(\vec{p},r)} d(\vec{p},\vec{q},r) g(\vec{p},\vec{q})}{\sum_{\vec{q}\in N(\vec{p},r)} d(\vec{p},\vec{q},r)} \tag{2}$$

where r is a neighborhood size threshold. The distance function $d(\vec{p},\vec{q},r) = (r - \|\vec{p}-\vec{q}\|)^2$ gives more weights to the particles near the center. We have applied our method to multiple particle data sets, including a chemical simulation data set composed of 100 time steps with around 2,000,000 three-dimensional particles in each time step. Figure 1 shows the architecture of our particle latent autoencoder and Figure 2 shows examples of reconstructed particles from particle latents in a time sequence. Experimental results showed that our approach can adapt to large particle data with multi-dimensional physical attributes and capture salient features by the latent representations. Our method provides an efficient way of using particle latents to extract related regions for tracking and further analysis. The particle latents are also used as the representations of particles for subsequent data aggregation and reduction, as explained later in the proposal.

2.1.2 Displacement Fields for Particle Positions

One factor that hinders the effectiveness of compressing particle data and extracting spatially coherent region is the absence of meshes to connect the discrete particles. This makes it more difficult to group particles in its spatial proximity without explicitly searching the domain. The absence of meshes also makes it cumbersome to form multiscale particle clusters, aggregate the physical properties, and sample particles from local regions adaptively. In contrast to the meshless particle data, there has been a rich set of techniques that can perform multiscale filtering, analysis, and data reduction for data defined on regular meshes, but unfortunately very few of which can be used for particle data.

In this research, an important stage of our particle data reduction pipeline is to perform a transformation of particle positions to address the aforementioned challenges. Specifically, our goal is to transform particles to a representation that facilitates efficient adaptive multiscale data analysis, reduction, and reconstruction. To achieve the goal, we seek a representation that can organize the particles losslessly in a structured way. We propose to transform the collection of particle positions to a regular Cartesian mesh $\mathbb{D}$ where each particle is associated with one and only one grid point that is the closest to it. With such a mesh $\mathbb{D}$, for every particle p and its associated grid point q, we store the displacement vector $p - q$ at q which is effectively a displacement vector of particle p from q. This effectively transforms a particle data set to a vector field defined on a regular Cartesian mesh $\mathbb{D}$.

Given a regular Cartesian mesh with its cell size equal to h, it can be easily seen that the length of any displacement vector stored in the mesh will not be larger than half of the cell length times $\sqrt{2}$, i.e., $|q-p| \leq \sqrt{2}h/2$, where p is a particle and q is the particle's associated grid point. To make sure each grid point has at most one particle associated with it, the resolution of the mesh $\mathbb{D}$ should have its cell size smaller than the distance of the closest particle pair in the data set divided by $\sqrt{2}$, denoted as c. To make sure that the mesh is not unnecessarily dense, we should choose the cell size of the mesh to be as large as possible but smaller than c. For grid points that have no particles in their local neighborhood, we can simply store a zero vector. The particle latents that represent the physical attributes can also be stored in the corresponding grid points at mesh $\mathbb{D}$. To resolve individual particles, the mesh of the displacement vectors will be dense, so this representation may be larger then the original particle data. However, the regular structure of the mesh is used to perform multiscale analysis and adaptive refinement as will be shown in the following sections, and hence there is no need to store such a dense mesh permanently. We also note that the constraint that each grid point have at most one particle can be relaxed. That is, a grid point can be associated with multiple particles, and the displacement vectors and the particle associated latents can be aggregated to perform multiscale analysis and data reduction, as explained in the next section.

2.2 Multiscale Analysis, Reduction, and Super-Resolution

The aforementioned transformed data lend themselves well to multiscale analysis and data reduction because the regular mesh makes it easier to discover structural coherence in local regions, remove redundancy, and preserve salient features that are otherwise difficult to find. The regularity of the mesh also makes adaptive coarsening and refinement of the features much more efficient. Furthermore, the transformed format will allow us to take advantage of results from recent advances of deep learning research that utilize convolutional neural networks which require regular grid meshes for feature extraction, classification, and quality enhancement. Below we describe our approaches in detail.

2.2.1 Distribution-based Multiscale Aggregation of Particle Attributes

While the regular grid displacement field for particle positions and the high dimensional latent representation for particle physical attributes allow us to perform efficient search of coherent structures, it is not storage space efficient because scientific features are often not evenly distributed in space. The regular grid representation, however, makes it simple to perform adaptive coarsening and refinement of data based on the locations of particles and the existence of data coherence in multiscale local neighborhoods. In this research, we will first perform a hierarchical spatial partitioning of the regular mesh, and then aggregate both the positions and the physical attributes. Data reduction can be achieved by data aggregation in the adaptively coarsened spatial local regions.

To perform spatial partitioning of the mesh to form a multiscale hierarchy, we will perform a recursive subdivision of the space based on pre-selected criteria with respect to the particle positions and properties. This multiscale hierarchy can be represented as a tree where the root of the tree represents the entire domain. Starting from the root, we check if the criteria for terminating further subdivision have been met, for example, if the number of particles, the spread of particle positions, the variance of particles' physical attributes, or the error associated with the aggregation function in use, is less than the pre-set thresholds. If not, the current regions in question will be split and the particles will be subdivided into the split subspaces. To allow the subdivision scheme to work with the spatial super-resolution work we propose later in Section 2.2.2, we will use the octree [47] as our hierarchical data structure.

Once the space partitioning is completed, based on the quality threshold, each leaf node in the hierarchical data structure will receive one or multiple particles with their attributes in the form of displacement vectors and particle latents. In the case that multiple particles are assigned to a leaf node, data reduction will be performed by aggregating the positional information (displacement vectors) and physical attributes (particle latents). With an appropriate aggregation function, the original data will be transformed to a multiscale representation with much smaller sizes.

The data in the leaf nodes of the octree for aggregation are the displacement vectors and particle latents, both of which can be considered as multi-dimensional particle attributes. Although we want to store the aggregated displacements and latents separately so that particle positions and physical attributes can be sampled and reconstructed separately for analysis, it is possible to have a unified approach for aggregating both. We propose to use multi-dimensional distribution models to aggregate and represent them in each leaf node. A distribution representation for multi-dimensional data can achieve a good trade-off between storage size and data quality and supports particle-based data analysis and visualization in the post data processing stage. Kernel Density Estimation (KDE) [65] is one of the popular non-parametric distribution-based models, but it incurs higher storage and computational costs. The histogram is another common representation that can be computed quickly, but it consumes more storage when representing a multi-dimensional distribution. For parametric distributions, Gaussian-based methods are commonly used, and can compactly represent a distribution because only a few parameters such as means and covariance have to be stored. When the data distribution is complex, a mixture of Gaussians called Gaussian Mixture Models (GMMs) is more suitable. A GMM uses K Gaussian models to approximate a data distribution. It has high storage efficiency and does not need prior assumptions about the distribution. Some related studies have pointed out that the use of GMMs is an efficient statistical data summarization method [17, 44]. In this research, we propose to use GMMs to approximate the distribution of the particle displacements and latent attributes in each leaf node. Equation 3 shows the formula of a GMM.

$$p_\theta(x) = \sum_{i=1}^{K} \omega_i * \mathcal{N}(\mu_i, \sigma_i) \tag{3}$$

where K is the number of Gaussian components in the GMM and θ represents GMM parameters which consist of weights, means and standards deviations of Gaussian components; ω_i, μ_i and σ_i are the weight, mean and standard deviation for the i^{th} Gaussian component, respectively. The sum of weights must be equal to 1.

To estimate the parameters of a GMM, the Expectation Maximization (EM) [16] algorithm is usually used. The EM algorithm interactively estimates the parameters of a model which can maximize the likelihood of the given data samples. We will apply the EM algorithm on particles of each leaf node of the hierarchy to calculate the parameters of a multi-dimensional GMM. Finally, we store the GMMs to represent the particle position and physical attributes. The number of components in GMMs is an important parameter that influences the storage size and quality when using GMMs. In this research, we will study how to select the number of Gaussian components for both particle displacement vectors and latents.

2.2.2 Spatial Super-Resolution with Hierarchial GANs

To support scientists performing post-hoc analysis with the highest possible data quality, it is important to reconstruct the particle data from the reduced multiscale representation described in the previous sections. Specifically, this includes reconstructing the displacement field and the aggregated latent representations to satisfy a user desired quality. From the reduced hierarchical mesh, reconstructing the displacement vectors to a higher resolution mesh allows us to recover the positions of individual particles or the mean positions and the covariances of particle clusters. Reconstructing higher resolution particle latents allows us to obtain the attributes of the corresponding particle(s) for further analysis. Since our displacement and particle latent fields can be considered as a multi-channel scalar/vector field, in this research we will develop a unified model for reconstruction using a deep learning based super-resolution approach. Below we describe our proposed approach in detail.

For data reconstruction, interpolation techniques like linear or bicubic interpolation have been commonly used to fill in the discarded data post-hoc, but can lead to overly smooth results and missing high-frequency features important for analysis. Due to the recent advancements of machine learning, super-resolution (SR) methods applied to scientific data have seen better feature reconstruction than interpolation methods [89, 28, 25, 27]. Using neural networks (NNs) for non-error-bounded lossy compression has been studied extensively for image compression with results exceeding JPEG and JPEG2000 [1, 75, 5, 74, 72, 71]. A major benefit of machine learning for data reconstruction is its ability to preserve global features well due to the learned statistics across large spatial and temporal domains. Currently, the state of the art in ML-based super-resolution work assumes a single network trained for reconstructing the data by a fixed ratio, such as $N\times$ scale-up. However, this fixed network will only offer two compression options - either the data is downscaled or it is not by $N\times$. Since our reduced representations in the displacement and latent fields are defined in a hierarchical mesh, there is a need for research to allow more flexible data reduction and reconstruction. This will require to train a hierarchy of networks G to allow $2\times, 4\times, ..., 2^{|G|}\times$ super-resolution in each spatial dimension. Because particle data tend to have uneven distribution of features in space, regions with simple features can allow larger compression rate and hence still retain the quality even with a larger scale super-resolution. On the other hand, regions that contain complex features that are harder to retain with aggressive reduction should be allowed to reduce the reduction rate. Overall, this will allow higher compression ratios for the same error bounds.

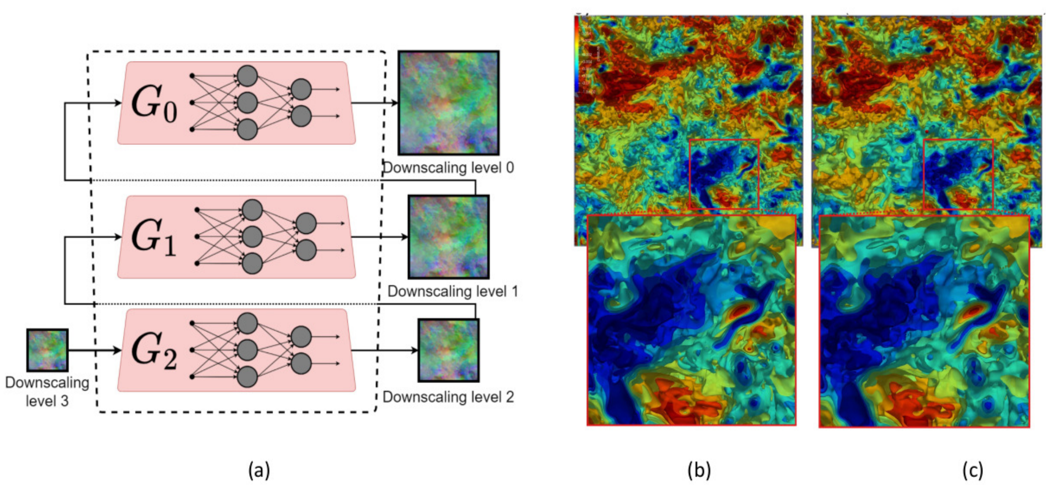

Figure 3: (a) An hierarchy of neural networks trained to perform $2\times$, $4\times$, or $8\times$ spatial super resolution. (b) the ground truth (c) the super-resolution result from the hierarchical neural network with 500 times compression.

In this research, we plan to use a hierarchy of super-resolution networks to facilitate error-controlled data reduction by allowing multiple scaling factors. We propose to use a hierarchy of identical generative adversarial networks (GANs), with each one trained to perform $2\times$ super-resolution in each spatial dimension between downscaling levels. We define downscaling level 0 as full resolution and downscaling level i is data that are downscaled by a factor of 2^i in each dimension. Together, they form a hierarchy of networks G that supports up to $2^{|G|}\times$ super-resolution. An example of a super-resolution hierarchy is shown in Figure 3(a). In our preliminary work [86], we use a super-resolution generator architecture based on an enhanced super-resolution generative adversarial network (ESRGAN) [66], a state-of-the-art GAN for up-scaling single images from a low resolution input to a high resolution output. Each generator is coupled with a patch discriminator, which classifies overlapping patches as real or fake. This network is designed to be relatively shallow with a small receptive field. The reason for us to choose GAN as a starting point of our research are twofold. First, they have shown to be powerful for reconstructing important high-frequency features for both images as well as scientific data. Second, we expect that output from a GAN, specifically Wasserstein GANs (WGANs) [4], will more accurately match the distribution of ground truth data, which is important when using multiple super resolution generators from the SR hierarchy in tandem. Figure 3 (b) and (c) show the ground truth and our super-resolution result with the proposed hierarchical GANs that can achieve more than 500 times of compression.

2.2.3 Temporal Super-Resolution with Motion Estimation

Almost all scientific data are time-varying where features evolve over time. To capture the dynamics of the natural phenomena, simulations often need to generate data in a long time sequence and as a result, the size of simulation output can easily overwhelm the resources available for I/O, network, and analysis. To address this challenge, in this research we will develop advanced schemes for time-varying particle data reduction and reconstruction, which will be combined with the spatial data super-resolution methods described in Section 2.2.2 to achieve effective spatiotemporal data reduction and reconstruction. Our focus will be on performing temporal super-resolution to recover data reduced by temporal subsampling, a common practice for time-varying data. As described previously, in our research spatial subsampling is achieved via multi-scale hierarchical data reduction. Temporal subsampling, on the other hand, means keeping only a subset of the spatially reduced data in the time sequence. To reconstruct the original particle data so as to support the scientists' analysis need, temporal super-resolution is performed to reconstruct low resolution spatial data at a higher temporal resolution, which are then sent to the spatial super-resolution module for a final spatiotemporal super-resolution. Below we describe our approach in detail.

Although there are existing works that utilize Long Short Term Memory (LSTM) [33] for temporal super-resolution, those methods mostly focus on updating the values at fixed grid locations over time from the discovered temporal patterns and thus can be considered as Eulerian approaches. For particle data, since there is an evolution of the locations and attributes both in space and time, it will allow more effective preservation of particle features by tracking the movement of particles when performing data reduction and reconstruction. Between adjacent time steps, such movement can be modeled as a motion field which can be learned by motion estimation considering the spatial and temporal coherence of the displacement vectors and particle latents. Given a time sequence $[i, i+1, ..., i+k]$, assuming the data are reduced by only saving the two end time steps V_i and V_{i+k}, a common practice used currently, our goal is to develop a deep learning based model to recover the missing time steps in between. We will perform temporal super resolution with motion estimation to generate the intermediate frames from both V_i and V_{i+k} bidirectionally. Our proposed method comprises of two components: motion estimation and fusion.

Deformation with motion estimation: Inspired by video frame interpolation works [6, 37], where optical flow $\mathbf{f}$ is estimated to determine the motion vector for each pixel in the frame, we focus on the dynamic changes along time between consecutive frames and aim to estimate the motion field. Let $V_t(x)$ denotes a time-varying field, where x denotes the spatial coordinates (2 or 3 dimensions) and t denotes time. In this way, the intermediate field can be estimated by aligning the previous frame with the motion to the current step, and then adding the residual. Likewise it can also be generated from the later frame. We can write the forward and backward alignments as:

$$\begin{aligned}\hat{V}_t^f(x) &= V_{t-1}(x - F_{t-1}(x)) + \Delta R_f \\ \hat{V}_t^b(x) &= V_{t+1}(x + F_t(x)) + \Delta R_f\end{aligned} \tag{4}$$

where $F_t(x)$ is the motion field.

To reconstruct the time sequence, we need the forward and backward motion field $F(x)$ at every time step i in between the two end frames, and the residual ΔR. However, since $F(x)$ is not available, we will use the motion field at two key frames i and $i+k$, denoted as $F_{i\to i+k}$ and $F_{i+k\to i}$, to approximate the motion fields we need: $\hat{F}_{i\to t}$ and $\hat{F}_{t\to i+k}$. In our preliminary work [3], we employ the U-Net architecture [64] as our motion estimation network. Also, we use a rectify block to learn the residual ΔR between the deformed vector fields and the ground truth field. Results have shown that we are able to recover the motion field effectively with U-Net.

Fusion After obtaining the forward and backward deformed frames $\hat{V}_t^f(x)$ and $\hat{V}_t^b(x)$, we introduce the modulation maps M_i, M_{i+1}, ..., M_{i+k}, where $M_i(x) \in [0,1]^2$ to denote the vector value from either forward and backward deformed frame contributing to the final result since generally averaging will produce blurry frame with artifacts. The modulation maps vary across the spatial dimension and will be trained in the same U-Net architecture taking $\hat{V}_t^f(x)$ and $\hat{V}_t^b(x)$ as input. Considering the temporal distance and the modulation map, the blended frame is

$$\hat{V}_t^{LR}(x) = (1 - \Delta t)M_t(x)[0] \cdot \hat{V}_t^f(x) + \Delta t M_t(x)[1] \cdot \hat{V}_t^b(x) \tag{5}$$

Using our model, we can increase the temporal sampling rate with reconstructed intermediate frames. These frames are still at a lower spatial resolution, denoted as $\hat{V}_t^{LR}(x)$ as we do not attemp to increase the spatial resolution here. Instead, the output from our temporal super-resolution model will be used for spatial super-resolution described in Section 2.2.2 to generate the final high-resolution frames.

2.3 Uncertainty Aware Particle Reconstruction

2.3.1 Error-controlled Particle Reconstruction and Sampling

When the scientists are ready to perform post-hoc analysis, it is necessary to reconstruct both the particle positions and physical attributes from the reduced representations. In the previous sections we described the process of data reduction via representation transformation, multiscale analysis, and spatial and temporal super-resolution. The output from the super-resolution components provides high quality displacement fields and aggregated latent representations for particle positions and physical attributes. Then, to allow scientists to visualize and analyze the actual particles in their original format, the final stage of our pipeline is to perform particle reconstruction and sampling. To allow a flexible and well-informed exploration of the resulting particle data, there are two important considerations. One is to recreate the particles from the transformed data, and the other is to present to scientists the uncertainty associated with the reconstructed data. Below we describe our research in detail.

As described previously, our reduced data are stored in a hierarchical data structure, octree, which allows for multiscale data reduction and reconstruction. Each leaf node of the octree stores the displacement vectors which describe the positions of individual particles or particle clusters, and the particles' physical attributes in the form of particle latents. The reduced data are represented as multi-dimensional distribution functions, for example as Gaussian mixture models, for both the displacement vectors and the particle latents. Since each octree node represents a spatial subdivision at a particular resolution, it also records the reconstruction error if the data are reconstructed at that resolution. In our data reduction pipeline, there will be two types of errors that will be stored in each octree nodes. One is the error associated with data aggregation, and the other is the error produced by our super-resolution algorithms using the deep learning models. For data aggregation error, in the case that Gaussian mixture models are used, we will calculate the the Akaike Information Criterion (AIC) as an estimator [56] as below:

$$AIC = 2k - 2\ln(\mathcal{L}(\theta|x)) \tag{6}$$

where k is the number of parameters in the GMM, and $\mathcal{L}(\theta|x))$ is the value of the maximum likelihood function. A lower AIC value means a better modeling quality. In this research, we will identify appropriate aggregation error functions for our reduced particle formats. As for the error caused by deep learning based super-resolution, the error is calculated based on the loss functions used for training. The errors are computed by applying the model to scale up the data from the resolution represented by the octree cells and compared with the ground truth.

With the errors stored in the octree nodes, the scientists can make an informed decision about what level of details the particles should be reconstructed to. Then, the final stage of particle reconstruction is to perform particle sampling from the high-dimensional distribution functions that describe the displacement vectors and particle latents. Essentially, the sampling can be considered as a two-step process. First, we will reconstruct particles positions by sampling from the aggregated displacement distributions stored in the octree node. When GMMs are used, sampling methods such as Box-Muller [11] can be used. Sampling from the distribution functions has a unique benefit that the probability $p(x)$ for a particle p to be found at a location $\mathbf{x}$ can be quantified. Then, from the sample particle location x, we can estimate the physical attributes of the particle in the form of particle latent h at x using the Bayes rule:

$$p(h|x) \approx p(x|h)p(h) \tag{7}$$

where $p(h)$ is the distribution function for the particle latent stored in the octree node, and $p(x|h)$ is the probability of x conditioned on h, which can be constructed by dividing the latent h into several subgroups and constructing a GMM for each of the latent subgroups. With equation 7, it becomes possible to quantify the uncertainty of the reconstructed particle positions and their attributes. This information offers many possibilities for developing uncertainty visualization algorithms for the reconstructed data which will be a focus of our research.

2.3.2 Information Analysis of DNN models

To validate the deep neural networks (DNN) models proposed above, it is important to analyze how these models preserve and distill information from the original data sets. Furthermore, it is important to support the ability to interpret, diagnose, and debug them. To achieve the goal, we will analyze the information flow across the DNN models via information theory. Originally developed by Claude E. Shannon, information theory is focused on understanding the limitations of a communication channel. The theory defines a message's information content as entropy, which characterizes the number of choices in the message that can be sent. With the definition of entropy, the capacity of a communication channel can be characterized even with the presence of distortion. Since the advent of information theory, it has been shown that the concept is so universal that the term "information" can be associated with many different semantics. In the context of this research, a DNN model can be considered as a communication channel that distills and transmits the input's most salient information to the output. For example, when the model is used to perform convolution, the DNN will extract salient information from the input and ignore irrelevant and redundant information, to identify the more important features. From this point of view, it is apparent that the entropy should progressively decrease as the input is traveling through different layers of the network, with the output unambiguously inferring the salient features of the input. The mutual information between the input and output, and between consecutive layers, can be used to characterize the quality of the preserved features, as well as to quantify the uncertainties of the model.

In this research, our goal is to understand the behavior of DNNs from (1) the diversity of the input scientific data as training or test data to the model, (2) the quality and uncertainty of the output reconstructed data or features from the model, (3) the information captured by each DNN layer and between the layers, (4) the information preserved in the feature maps or channels, and (5) the function of neurons. We will use entropy to quantify the information loss in each of the DNN entities and use the mutual information to compute how information is transmitted and refined through the network layers. To analyze model quality, we will develop a data model in the form of multidimensional arrays to describe the information generated during training and testing of the DNN models, such that each dimension represents one of the items listed in (1)-(5) described above.

We will also quantify the information in the original particle clouds to analyze how information is preserved in the latent vectors. Similar measures will be compared to the spatiotemporal super-resolution models. It is noteworthy that in the field of visualization, information theory has been widely applied to different problems [79, 66], such as the selection of camera locations [10], isosurfaces [83], level of detail [78], and placement of streamline seeds [88], all done by PI Shen and his associates in the past decade. In addition, PI Shen previously led an ASCR-sponsored Scientific Data Management and Analysis at Extreme Scale project that developed an information-theoretic framework for visualizing large-scale scientific data sets. Their research resulted in methods to identify salient variables [45], placement of data [82], and important values given a variable [18], just to name a few. In this project, we will expand the framework to include DNN analysis and validation. We anticipate that our research will have potential to benefit and generalize to a wider range of DNN applications.

3 Application Scenarios and Use Cases

3.1 Smooth Particle Hydrodynamics

The Lagrangian nature of SPH makes it ideally suited to model applications that, over time, span a wide range of length scales. One such application lies in astrophysical transients: supernovae, gamma-ray bursts, and the ejecta from neutron star mergers. These astrophysical transients are the dominant source of many of the elements heavier than Lithium in the universe and understanding the nature of these explosions takes us one step closer to understanding how the Earth and mankind formed. For example, the so-called “kilonova” ejecta from neutron star mergers has become one of the hottest topics in astrophysics because of the detection of these events with gravitational waves [2] and the proof this detection provided that neutron star mergers could produce the bulk of the heavy elements (created through the rapid capture of neutrons) in the universe. LANL scientists have actively been modeling the ejecta of these astrophysical transients, following the ejecta from their source at 10-100 km to their ultimate “remnant” distribution in the Milky Way $(10^{1}0 - 10^{1}2\,\mathrm{km})$. Radioactive isotopes produced in these explosions decay, producing gamma-ray emission observable by existing and planned NASA missions like NuSTAR [29] and COSI [39].

Our goal is to find ways to reduce the storage constraints of the particle data in a way that it can be more easily compared to existing and upcoming NASA missions. We will incorporate known smoothing kernels in our analysis to capture the results of the simulations and provide accurate predictions for the gamma-ray emission. By making the data more easily accessible to astronomers proposing new missions, we can help design upcoming satellites to maximize their science.

3.2 Monte Carlo Radiation Transport

Many high-energy transient events in the universe—such as compact object mergers and the deaths of massive stars—produce a disk of material accreting onto a central object, such as a black hole. This accretion powers an ultra-relativistic jet of material that flows out the poles of the system at almost the speed of light [48]. A long standing open question is whether this jet is powered by electromagnetic processes [9] or by neutrino annihilation [19]. Recent simulations coupling a MC treatment of neutrinos to a fluid [50, 51] have the potential to answer this question. However, they require an up-sampling of neutrinos in the middle of the simulation and a fast reconstruction of the neutrino distribution function. We will apply the super-resolution techniques described here to this upsampling and enable, finally, a resolution to the problem of jet power.

In MC transport, particles represent statistical samples of a six-dimensional phase space density function. When appropriately normalized, the phase space density can be viewed as a probability distribution representing the probability of finding a physical particle such as a photon or neutrino. An important property of MC simulations is the conservation of probabilities. Each particle carries with it a statistical weight and these weights must sum appropriately, so that the MC particles represent meaningful samples of a probability distribution. We will evaluate our methods, as applied to Monte Carlo, by checking how faithfully they can reconstruct the phase space density, and how well they preserve unitarity—the fact that the density is a probability distribution.

3.3 Tracer Particles

In Eulerian (non-SPH) simulations of neutron star mergers and core-collapse supernovae, tracer particles are often used to extract the time-histories of microphysical quantities within a single fluid packet. This is required for computing, e.g., the nucleosynthesis of heavy elements in these events. Temporal super-resolution could be very valuable in this domain, as a way to compress and up-sample the time-history of this data. Moreover, simulations often end before all interesting dynamics have ceased. One interesting application we will explore is using neural networks to extrapolate the time-history of our dataset for use in calculating, e.g., nucleosynthetic yields. To evaluate how succesful our approach is, we will pass reconstructed tracer histories through an r-process nucleosynthesis network such as SkyNet [43] or PRISM [54, 15, 91, 69] and evaluate how well the yields match the original data. Tracer data has the unique property that, while it is particle data, it is tied to a grid. This makes it an ideal testing ground for our data reduction techniques—since we can compare the grid data to the tracer data at every stage of our analyses. When evaluating our techniques, we will compare how well the reconstructed tracer data matches the grid data.

4 Project Management

4.1 Staffing and coordination

Han-Wei Shen is a full professor at The Ohio State University. He has nearly three decades of experience in large-scale scientific data analysis and visualization research. He was the recipient of 2003 DOE Early Career Principle Investigator (ECPI) and 2004 NSF CAREER awards. He also served as the lead PI for two previous ASCR Scientific Data Management, Analysis, and Visualization projects, a co-PI for two ASCR-funded computer science projects, and a co-PI for four SciDAC projects since 2006, including the most recent SciDAC institute for Computer Science, Data, and Artificial Intelligence (RAPIDS2). In 2020, He was inducted into IEEE VGTC Visualization Academy, a prestigious honor given to the most accomplished visualization researchers internationally. He is currently an Associate Editor-in-Chief for IEEE Transactions on Visualization and Computer Graphics.

Ayan Biswas is a scientist at Los Alamos National Laboratory. He is currently leading the sampling-based data reduction efforts on the LANL side under the Exascale Computing Project (ECP) for ALPINE. He is also the co-PI for a LANL funded LDRD-DR project (total funding $5M over 3 years) that focuses on in-situ data modeling and statistical inference. He is a visualization and big data analysis expert with also considerable experience in information theoretic approaches, uncertainty quantification, and machine learning.

Chris Fryer is a scientist 6 at Los Alamos National Laboratory. He is co-developer of the SNSPH smooth particle hydrodynamics code. For his work modeling astrophysical phenomena (over 200 refereed papers, many using smooth particle hydrodynamics techniques and including the first 3-dimensional models of core-collapse supernovae), he was given the E.O. Lawrence award and named a fellow of the APS, the AAS, and LANL. He is currently the chair of the division of astrophysics of the APS.

Jonah Miller is a scientist at Los Alamos National Laboratory. He is the lead developer of `nubhlight`, a Monte Carlo radiation hydrodynamics code with tracer particles. `nubhlight` is currently the only code in the world capable of modeling the aftermath of a neutron star merger with all the relevant physics included. Miller is an expert in compact object astrophysics and high performance computing, with nearly a decade of experience in both fields. Codes co-developed by Miller have achieved extreme scaling by, for example, running on the entirety of the Blue Waters supercomputer.

During the project period, PI Shen will oversee the execution of the entire proposed research. He will also lead his research team to tackle the problems of particle data transformation, spatiotemporal super-resolution, and information analysis of DNN models. PI Biswas will lead the effort of particle data aggregation, multiscale analysis, and uncertainty-aware particle reconstruction and sampling. PIs Fryer and Miller will be responsible for preparing application test data, specifying application scenarios and use cases, and performing application-centric evaluation of our proposed research methods. Monthly teleconferences will be used to report progress, to present research ideas, and to prioritize execution plans. All the PIs and the postdoctoral and student researchers will participate. In addition, we will organize annual PI meetings. The kick-off meeting will be held at LANL so that the PIs can have an in-depth discussion and understanding of the application needs, followed by at OSU and LANL sites in the second and final years respectively. Student researchers from OSU will go to LANL during the summer of each year to work closely with the scientists. We will set up a website, code repository, wiki, and mailing list for this project. All source code and publications will be available to the public online (see Appendix 6 for details on data management plan).

FY 22				FY 23				FY 24			
Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Particle Latent Extraction (Sec. 2.1.1)				Displacement Aggregation (Sec. 2.1.1)		Spatial Super-Resolution (Sec. 2.2.2)			Uncertainty-aware Particle Reconstruction (Sec. 2.3.1)		
Displacement Field Computation (Sec. 2.1.2)			Hierarchical Decomposition (Sec. 2.2.1)			Latent Aggregation (Sec. 2.1.1)		Temporal Super-Resolution and Reconstruction (Sec. 2.2.3)			
			Information Analysis of DNN Models (Sec. 2.3.2)								
Application Data Preparation and Use Cases Specification (Sec. 3)							Algorithms and Applications Use Cases Evaluation (Sec. 3)				

Table 2: Project timeline at a glance.

4.2 Milestones and Deliverables

Table 2 lists the milestones of this project. Our year 1 effort will be focused on developing the proposed data transformation methods and also working closely with applications to prepare test data and define detailed application use cases. Then we will proceed to develop multiscale analysis techniques for data aggregation, reduction and super-resolution in year 2 and 3. Since DNN models will be developed, we will simultaneously perform information analysis for the DNN models. Finlly the focus of year 3 will be on uncertainty-aware particle reconstruction, and performing a thorough evaluation of our approaches for scientific applications and the use cases. We will disseminate our results to researchers in DOE ASCR and the broader community by creating stand alone libraries for data transformation, multiscale analysis, and uncertainty quantification. Specifically, we will release our deep learning models for particle latent transformation, spatial super-resolution with hierarchical GANs, and neural networks for motion estimation and temporal super-resolution. We will also create software for transforming particle positions to hierarchical displacement representations, the distribution-based particle aggregation such as Gausian Mixture Models, and the corresponding particle re-sampling modules. Finally, we will attempt to integrate our models with production visualization tools including ParaView and VTK. All software deliverables will be open source and publicly available for download. Details on code hosting, version controlling, and software licensing are in the data management plan.

The PIs will engage in activities beyond academic publication and software dissemination with the goal of benefiting the wider community of scientists and engineers. The PIs will organize tutorials on particle analysis and visualization methods at annual ACM/IEEE Supercomputing, and IEEE VGTC-sponsored visualization conferences such as the IEEE Visualization and IEEE Pacific Visualization conferences. Furthermore, the PIs will introduce the particle data reduction and multiscale analysis methods and software to a wider audience in domain-specific scientific conferences and meetings.

Appendix 1: Biographical Sketches

NAME: Han-Wei Shen

POSITION TITLE & INSTITUTION: Full Professor, The Ohio State University

A. PROFESSIONAL PREPARATION
(see PAPPG Chapter II.C.2.f.(i)(a))

INSTITUTION	LOCATION	MAJOR/AREA OF STUDY	DEGREE (if applicable)	YEAR (YYYY)
National Taiwan University	Taipei, Taiwan	Computer Science	BS	1988
Stony Brook University	Stony Brook, NY	Computer Science	MS	1992
University of Utah	Salt Lake City, UT	Computer Science	PhD	1998

B. APPOINTMENTS
(see PAPPG Chapter II.C.2.f.(i)(b))

From - To	Position Title, Organization and Location
1996-1999	Research Scientist, NASA Ames Research Center
1999-2005	Assistant Professor, The Ohio State University
2015-2012	Associate Professor, The Ohio State University
2012-Present	Full Professor, The Ohio State University

C. PRODUCTS
(see PAPPG Chapter II.C.2.f.(i)(c))

Products Most Closely Related to the Proposed Project

1, He W, Wang J, Guo H, Wang KC, Shen HW, Raj M, Nashed YSG, Peterka T. InSituNet: Deep Image Synthesis for Parameter Space Exploration of Ensemble Simulations. IEEE Trans Vis Comput Graph. 2020 Jan;26(1):23-33. PubMed PMID: 31425097.

2. Hazarika S, Li H, Wang KC, Shen HW, Chou CS. NNVA: Neural Network Assisted Visual Analysis of Yeast Cell Polarization Simulation. IEEE Trans Vis Comput Graph. 2020 Jan;26(1):34-44. PubMed PMID: 31425114.

3. Wang J, Gou L, Shen HW, Yang H. DQNViz: A Visual Analytics Approach to Understand Deep Q-Networks. IEEE Trans Vis Comput Graph. 2018 Sep 5;PubMed PMID: 30188823.

4. Wang J, Gou L, Yang H, Shen HW. GANViz: A Visual Analytics Approach to Understand the Adversarial Game. IEEE Trans Vis Comput Graph. 2018 Jun;24(6):1905-1917. PubMed PMID: 29723140.

5. Dutta S, Chen CM, Heinlein G, Shen HW, Chen JP. In Situ Distribution Guided Analysis and Visualization of Transonic Jet Engine Simulations. IEEE Trans Vis Comput Graph. 2017 Jan;23(1):811-820. PubMed PMID: 27875195.

Other Significant Products, Whether or Not Related to the Proposed Project

1. Biswas A, Lin G, Liu X, Shen HW. Visualization of Time-Varying Weather Ensembles across Multiple Resolutions. IEEE Trans Vis Comput Graph. 2017 Jan;23(1):841-850.
2. Dutta S, Shen HW. Distribution Driven Extraction and Tracking of Features for Time-varying Data Analysis. IEEE Trans Vis Comput Graph. 2016 Jan;22(1):837-46
3. Liu X, Shen HW. Association Analysis for Visual Exploration of Multivariate Scientific Data Sets. IEEE Trans Vis Comput Graph. 2016 Jan;22(1):955-64.
4. Chen CM, Dutta S, Liu X, Heinlein G, Shen HW, Chen JP. Visualization and Analysis of Rotating Stall for Transonic Jet Engine Simulation. IEEE Trans Vis Comput Graph. 2016, Jan;22(1):847-56.

D. SYNERGISTIC ACTIVITIES
(see PAPPG Chapter II.C.2.f.(i)(d))

1. Associate Editor-in-Chief, IEEE Transactions on Visualization and Computer Graphics 2019-Present
2. IEEE Visualization SciVis Paper co-Chair, 2013, 2014, 2020
3. Chair, IEEE SciVis Steering Committee, 2018 – 2020
4. IEEE Visualization Executive Committee, 2016 - 2020
5. IEEE Pacific Visualization Paper co-Chair 2009, 2010

Collaborators, Co-editors, and Graduate and Postdoctoral Advisors and Advisees of Han-Wei Shen

Last Name	First Name	Title	Institution
Ahrens	James	Senior Scientist	Los Alamos National Laboratory
Balaprakash	Prasanna	Scientist	Argonne National Laboratory
Childs	Hank	Professor	University of Oregon
Dutta	Soumya	Scientist	Los Alamos National Laboratory
Geveci	Berk	Senior Director of Scientific Computing	Kitware Inc.
Guo	Hanqi	Scientist	Argonne National Laboratory
Hansen	Chuck	Professor	University of Utah
Hazarika	Subhashis	Postdoctorate Researcher	Los Alamos National Laboratory
He	Wenbin	Scientist	Bosch Research
Johnson	Christopher	Professor	University of Utah
Kaufman	Arie	Professor	Stonybrook University
Klasky	Scott	Scientist	Oak Ridge National Laboratory
Liu	Xiaotong	Research Staff	IBM

Ma	Kwan-Liu	Professor	University of California Davis
Morozov	Dmitriy	Scientist	Lawrence Berkley National Laboratory
Mueller	Klaus	Professor	Stonybrook University
Nouanesengsy	Boonthanome	Scientist	Los Alamos National Laboratory
Oliker	Leonid	Scientist	Lawrence Berkley National Laboratory
Patchett	John	Scientist	Los Alamos National Laboratory
Peterka	Tom	Scientist	Argonne National Laboratory
Pugmire	David	Scientist	Oak Ridge National Laboratory
Ross	Rob	Scientist	Argonne National Laboratory
Wang	Chaoli	Associate Professor	University of Notre Dame
Wang	Junpeng	Scientist	VISA Research
Wu	John	Scientist	Lawrence Berkley National Laboratory
Woodring	Jonathan	Scientist	Los Alamos National Laboratory
Yoo	Shinjae	Scientist	Brookhaven National Laboratory

Identification of Potential Conflicts of Interest or Bias in Selection of Reviewers

N/A

NAME: Chris Fryer

POSITION TITLE & INSTITUTION: Scientist 6, LANL

A. PROFESSIONAL PREPARATION
(see PAPPG Chapter II.C.2.f.(i)(a))

INSTITUTION	LOCATION	MAJOR/AREA OF STUDY	DEGREE (if applicable)	YEAR (YYYY)
UC Berkeley	Berkeley, CA	Mathematics/Astronomy	B.A.	1992
U Arizona	Tucson, AZ	Astronomy	PhD	1996

B. APPOINTMENTS
(see PAPPG Chapter II.C.2.f.(i)(b))

From - To	Position Title, Organization and Location
1996-2000	Post-doc, UC Santa Cruz, Santa Cruz, CA
2000-2002	Feynman Fellow, Los Alamos National Laboratory, Los Alamos, NM
2002-present	Staff Scientist, Los Alamos National Laboratory, Los Alamos, NM
2003-	Adjunct Faculty, Physics Dept., University of Arizona, Tucson, AZ
2009-	Adjunct Faculty, Physics and Astronomy Dept., University of New Mexico, Albuquerque, NM
2017-	Adjunct Faculty, George Washington University

C. PRODUCTS

(see PAPPG Chapter II.C.2.f.(i)(c))

Products Most Closely Related to the Proposed Project

1) PI, LANL Kilonova effort with database of spectra and light-curves:
https://ccsweb.lanl.gov/astro/transient/transients_astro.html

2) Shock heating in early-time transient emission:
Fryer et al., "The Role of Inhomogeneities in Supernova Shock Breakout Emission", 2020, ApJ, 898 123

3) Ties to Transport Experiments
Fryer et al., "Designing radiation transport tests: Simulation-driven uncertainty-quantification of the COAX temperature diagnostic", 2020, HEDP, 3500738

Studies of compact remnant mass distributions:
4) Fryer & Kalogera, "Theoretical Black Hole Mass Distributions", 2001, ApJ, 554, 548;
5) Fryer et al., "Compact Remnant Mass Function: Dependence on the Explosion Mechanism and Metallicity", 2012, ApJ, 749, 91

Other Significant Products, Whether or Not Related to the Proposed Project

1) Developer of SNSPH (Smooth Particle Hydrodynamics Code): Fryer, Rockefeller, & Warren, "SNSPH: A Parallel Three-dimensional Smoothed Particle Radiation Hydrodynamics Code", 2006, ApJ, 643, 292
2) Developer in LANL's ASC program: e.g.: Fryer et al., "The Role of Inhomogeneities in Supernova Shock Breakout Emission", 2020, ApJ, 898, 123

D. SYNERGISTIC ACTIVITIES

(see PAPPG Chapter II.C.2.f.(i)(d))

chair, Division of Astrophysics, American Physical Society
Project Scientist, SIBEX satellite mission proposal
Science team lead, ASCENT

Chris Fryer
Collaborators and Co-editors:

Last Name	First Name	Title	Institution
Becerra	Laura	Researcher	ICRANet
Belczynski	Krzysztof	Professor	Polish Academy of Sciences
Benz	Willy	Professor	University of Berne
Brown	Peter	Research Scientist	Texas A&M University
Coffing	Shane	Scientist	Los Alamos National Laboratory
Doctor	Zoheyr	Postdoctoral Scholar	University of Oregon
Even	Wesley	Scientist	Los Alamos National Laboratory
Fitz-Axen	Margot	Student	University of Texas at Austin
Fontes	Christopher	Scientist	Los Alamos National Laboratory
Grefenstette	Brian	Scientist	California Institute of Technology
Hartmann	Dieter	Professor	Clemson University
Holz	Daniel	Professor	University of Chicago
Kalternborn	Alex	Professor	GWU
Kislat	Fabian	Professor	University of New Hampshire
Korobkin	Oleg	Scientist	Los Alamos National Laboratory
Kouveliotou	Chryssa	Professor	George Washington University
Lloyd-Ronning	Nicole	Scientist	Los Alamos National Laboratory
O'Shaughnessy	Richard	Professor	Rochester Institute of Technology
Rockefeller	Gabriel	Scientist	Los Alamos National Laboratory
Roming	Peter	Scientist	Southwest Research Institute
Rosswog	Stephan	Professor	Stockholm University

Ruffini	Remo	Director	International Centre for Relativistic Astrophysics Network
Ryan	Benjamin	Scientist	Los Alamos National Laboratory
Safi-Harb	Samar	Professor	University of Manitoba
Timmes	Francis	Professor	Arizona State University
Troja	Eleonora	Associate Research Scientist	University of Maryland
Wollaeger	Ryan	Scientist	Los Alamos National Laboratory
Young	Patrick	Professor	Arizona State University

Identification of Potential Conflicts of Interest or Bias in Selection of Reviewers

N/A

NAME: Ayan Biswas

POSITION TITLE & INSTITUTION: Scientist 2, Los Alamos National Laboratory

A. PROFESSIONAL PREPARATION
(see PAPPG Chapter II.C.2.f.(i)(a))

INSTITUTION	LOCATION	MAJOR/AREA OF STUDY	DEGREE (if applicable)	YEAR (YYYY)
Jadavpur University	Kolkata, India	Computer Science and Engineering	BS	2007
The Ohio State University	Columbus, OH	Computer Graphics and Data Visualization	MS	2015
The Ohio State University	Columbus, OH	Computer Graphics and Data Visualization	PhD	2016
Los Alamos National Laboratory	Los Alamos, NM	Postdoctoral Researcher, Data Science and Visualization		2017-2018

B. APPOINTMENTS
(see PAPPG Chapter II.C.2.f.(i)(b))

From - To	Position Title, Organization and Location
2018-present	Scientist, Los Alamos National Laboratory, Los Alamos, NM
2017-2018	Postdoctorate Researcher, Los Alamos National Laboratory, Los Alamos, NM
2012-2015 (summer)	Graduate Student Intern, Los Alamos National Laboratory, Los Alamos, NM
2011-2016	Graduate Research Assistant, Ohio State University, Columbus, OH
2010-2011	Graduate Teaching Assistant, Ohio State University, Columbus, OH
2007-2010	Software Engineer, STMicroelectronics, UP, India

C. PRODUCTS

(see [PAPPG Chapter II.C.2.f.(i)(c)])

Products Most Closely Related to the Proposed Project

Ayan Biswas , Soumya Dutta, Earl Lawrence, John Patchett, Jon C. Calhoun, James Ahrens, 2020, Probabilistic Data-Driven Sampling via Multi-Criteria Importance Analysis, IEEE Transactions on Visualization and Computer Graphics

Pascal Grosset, Christopher M. Biwer, Jesus Pulido, Arvind T. Mohan, Ayan Biswas, John Patchett, Terece L. Turton, David H. Rogers, Daniel Livescu, James Ahrens, 2020, Foresight: Analysis That Matters for Data Reduction, Supercomputing 2020

Soumya Dutta, Ayan Biswas, James Ahrens, 2019, Multivariate Pointwise Information-driven Data Sampling and Visualization, Entropy journal.

Ayan Biswas, Soumya Dutta, Jesus Pulido, and James Ahrens, 2018, In Situ Data-Driven Adaptive Sampling for Large-scale Simulation Data Summarization, In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization (ISAV 2018)

Matthew Larsen, Amy Woods, Nicole Marsaglia, Ayan Biswas, Soumya Dutta, Cyrus Harrison, and Hank Childs, 2018, "A Flexible System for In Situ Triggers", (**Best Paper**), ISAV 2018

Other Significant Products, Whether or Not Related to the Proposed Project

Ayan Biswas, Christopher M. Biwer, David J. Walters, James Ahrens, Devin Francom, Earl Lawrence, Richard L. Sandberg, D. Anthony Fredenburg, and Cynthia Bolme, 2018, An Interactive Exploration Tool for High-Dimensional Datasets: A Shock Physics Case Study, Computing in Science and Engineering (CISE) 2018

David J Walters, Ayan Biswas, Earl C Lawrence, Devin C Francom, Darby J Luscher, D Anthony Fredenburg, Kelly R Moran, Christine M Sweeney, Richard L Sandberg, James P Ahrens, CA Bolme , 2018, Bayesian calibration of strength parameters using hydrocode simulations of symmetric impact shock experiments of Al-5083, Journal of Applied Physics, Volume 124, 2018

D. SYNERGISTIC ACTIVITIES

(see [PAPPG Chapter II.C.2.f.(i)(d)])

Member of International Program Committee (IPC) for IEEE Vis 2018-2021 (SciVis Short Papers), ISVC 2018-2021

Video Previews and Fast-Forward Chair for IEEE Vis 2018-2019

Meetups Chair for IEEE Vis 2020-2021

Ayan Biswas
Collaborators and Co-editors:

Last Name	First Name	Title	Institution
Ahrens	James	Senior Scientist	Los Alamos National Laboratory
Almgren	Ann	Senior Scientist	Lawrence Berkeley National Laboratory
Banesh	Divya	Postdoctorate Researcher	Los Alamos National Laboratory
Barber	Jon	Scientist	Los Alamos National Laboratory
Biwer	Christopher	Scientist	Los Alamos National Laboratory
Bolme	Cynthia	Scientist	Los Alamos National Laboratory
Brislawn	Christopher	Scientist	Los Alamos National Laboratory
Bryan	Chris	Professor	Arizona State University
Bujack	Roxana	Scientist	Los Alamos National Laboratory
Calhoun	Jon	Professor	Clemson University
Casleton	Emily	Scientist	Los Alamos National Laboratory
Childs	Hank	Professor	University of Oregon
Dutta	Soumya	Scientist	Los Alamos National Laboratory
Francom	Devin	Scientist	Los Alamos National Laboratory
Fulp	Megan	Student	Clemson University
Geveci	Berk	Senior Director of Scientific Computing	Kitware Inc.
Grosset	Pascal	Scientist	Los Alamos National Laboratory
Guo	Hanqi	Scientist	Argonne National Laboratory
Harrison	Cyrus	Scientist	Lawrence Livermore National Laboratory
Hazarika	Subhashis	Postdoctorate Researcher	Los Alamos National Laboratory
He	Wenbin	Scientist	Bosch Research

Johnson	Christopher	Professor	University of Utah
Keefe	Daniel	Professor	University of Missesota
Larson	Matthew	Scientist	Lawrence Livermore National Laboratory
Lawrence	Earl	Scientist	Los Alamos National Laboratory
Liu	Qun	Student	Louisiana State University
Liu	Xiaotong	Research Staff	IBM
Livescu	Daniel	Scientist	Los Alamos National Laboratory
Luscher	Darby	Scientist	Los Alamos National Laboratory
Ma	Kwan-Liu	Professor	University of California Davis
Mishra	Aditi	Student	Arizona State University
Mohan	Arvind	Postdoctorate researcher	Los Alamos National Laboratory
Moran	Kelly	Scientist	Los Alamos National Laboratory
Morozov	Dmitriy	Scientist	Lawrence Berkley National Laboratory
Musser	Jordan	Research Scientist	National Energy Technology Laboratory
Oliker	Leonid	Scientist	Argonne National Laboratory
Orban	Daniel	Student	University of Missesota
Patchett	John	Scientist	Los Alamos National Laboratory
Peterka	Tom	Scientist	Argonne National Laboratory
Pulido	Jesus	Scientist	Los Alamos National Laboratory
Ramos	Kyle	Scientist	Los Alamos National Laboratory
Rogers	David	Scientist	Los Alamos National Laboratory

Ross	Rob	Scientist	Argonne National Laboratory
Saavedra	Ramon	Scientist	Los Alamos National Laboratory
Sandberg	Richard	Professor	Brigham Young University
Sjue	Sky	Scientist	Los Alamos National Laboratory
Sweeney	Christine	Scientist	Los Alamos National Laboratory
Turton	Terece	Scientist	Los Alamos National Laboratory
Urban	Nathan	Scientist	Brookhaven National laboratory
Vogel	Sven	Scientist	Los Alamos National Laboratory
Walters	David	Scientist	Los Alamos National Laboratory
Wang	Chaoli	Associate Professor	University of Notre Dame
Wang	Junpeng	Scientist	VISA Research
Wolfram	Philip	Scientist	Los Alamos National Laboratory
Woodring	Jonathan	Scientist	Los Alamos National Laboratory

Identification of Potential Conflicts of Interest or Bias in Selection of Reviewers

N/A

NAME: Jonah Miller

POSITION TITLE & INSTITUTION: Scientist 2, Los Alamos National Laboratory

A. PROFESSIONAL PREPARATION
(see PAPPG Chapter II.C.2.f.(i)(a))

INSTITUTION	LOCATION	MAJOR/AREA OF STUDY	DEGREE (if applicable)	YEAR (YYYY)
CU Boulder	Boulder, CO	Physics/Mathematics	B.A.	2013
U of Guelph	Guelph, ON, CA	Physics	PhD	2017

B. APPOINTMENTS
(see PAPPG Chapter II.C.2.f.(i)(b))

From - To	Position Title, Organization and Location
2017-2019	CNLS Fellow, Los Alamos National Laboratory, Los Alamos, NM
2019-present	Staff Scientist, Los Alamos National Laboratory, Los Alamos, NM

C. PRODUCTS
(see PAPPG Chapter II.C.2.f.(i)(c))
Products Most Closely Related to the Proposed Project

Co-I LANL Kilonova effort. Lead on modeling of post-merger disks.
1) Miller et al. "Full transport model of GW170817-like disk produces a blue kilonova," 2019, PRD 100, 023008.
2) Miller et al. "Full transport general relativistic radiation magnetohydrodynamics for nucleosynthesis in collapsars," 2020, ApJ 202, 66.

Principle developer of general relativistic neutrino radiation magnetohydrodynamics code, nubhlight. The code utilizes both Monte Carlo radiation transport and tracer particles.
3) https://github.com/lanl/nubhlight
4) Miller et al. "nubhlight: radiation GRMHD for neutrino-driven accretion flows," 2019, ApJS 241, 30.

Monte Carlo transport for core-collapse supernova light curves and spectra
5) Curtis, ..., Miller et al. "Core-collapse supernovae: from neutrino-driven 1D explosions to light curves and spectra," 2020, submitted to ApJ. ArXiv:2008.05498

Other Significant Products, Whether or Not Related to the Proposed Project

Discontinuous Galerkin methods for high performance computing and numerical relativity
1) Miller and Schnetter, 2017, CQG, 34, 1.
2) Kidder, ..., Miller, et al. 2017, JCOMP, 335, 84.

Developer in LANL's ASC program, and contributor to many LANL-funded open-source projects:
3) https://github.com/lanl/parthenon
4) https://github.com/lanl/singularity-eos
5) https://github.com/lanl/singularity-opac

D. SYNERGISTIC ACTIVITIES
(see PAPPG Chapter II.C.2.f.(i)(d))

Member of American Astronomical Society
Participant in peer review
Volunteer speaker at Los Alamos Nature Center

Jonah Miller
Collaborators and Co-editors:

Last Name	First Name	Title	Institution
Allen	Gabrielle	Professor	University of Illinois Urbana-Champaign
Brockhauser	Sandor	Professor	European XFEL GmbH
Burrows	Adam	Professor	Princeton University
Carlson	Joseph	Scientist	Los Alamos National Laboratory
Chase	Eve	Graduate Research Associate	Los Alamos National Laboratory
Cooperman	Joshua	Lecturer	Bucknell University
Couture	Aaron	Scientist	Los Alamos National Laboratory
Curtis	Sanjana	Postdoctoral Associate	University of Amsterdam
Deppe	Nils	Graduate Research Associate	Cornell University
Diener	Peter	Professor	Louisiana State University
Dolence	Joshua	Scientist	Los Alamos National Laboratory
Ebinger	Kevin	Postdoctoral Associate	Technische Universitat Darmstadt
Etienne	Zachariah	Professor	West Virginia University
Even	Wesley	Scientist	Los Alamos National Laboratory
Fangohr	Hans	Professor	University of Southampton
Fields	Carla	Graduate Research Associate	Michigan State University and Arizona State University
Fields	Scott	Professor	University of Massachusetts Amherst
Fontes	Christopher	Scientist	Los Alamos National Laboratory
Foucart	Francois	Professor	University of New Hampshire
Francois	Hebert	Engineer	SpaceX
Frohlich	Carla	Professor	North Carolina State university
Fryer	Christopher	Scientist	Los Alamos National Laboratory

Gandolfi	Stefano	Scientist	Los Alamos National Laboratory
George	Daniel	Scientist	Google X
Haas	Roland	Scientist	National Center for Supercomputing Applications
Huerta	Eliu	Professor	University of Illinois Urbana-Champaign
Hungerford	Aimee	Scientist	Los Alamos National Laboratory
Kaltenborn	Mark	Graduate Research Associate	George Washington University
Katz	Daniel	Professor	University of Illinois Urbana-Champaign
Kidder	Lawrence	Professor	Cornell University
Korobkin	Oleg	Scientist	Los Alamos National Laboratory
Lee	Kyle	Graduate Research Associate	Stony Brook University
Lim	Hyun	Postdoctoral Associate	Los Alamos National Laboratory
Lippuner	Jonas	Scientist	Los Alamos National Laboratory
Lloyd-Ronning	Nicole	Scientist	Los Alamos National Laboratory
Loiseau	Julien	Scientist	Los Alamos National Laboratory
Martin	Phillip	Undergraduate Research Associate	Los Alamos National Laboratory
Misch	Wendell	Postdoctoral Associate	Los Alamos National Laboratory
Mottola	Emil	Scientist	Los Alamos National Laboratory
Mumpower	Matthew	Scientist	Los Alamos National Laboratory
Niemeyer	Kyle	Professor	Oregon State University
Ott	Christian	Professor	California Institute of Technology
Piotrowska	Joanna	Graduate Research Associate	University of Cambridge
Prasad	Neelima	Undergraduate Research Associate	Los Alamos National Laboratory

Queiroz	Francisco	Professor	Tecgraf Institute
Rosswog	Stephan	Professor	Stockholm University
Ryan	Benjamin	Scientist	Los Alamos National Laboratory
Salvesen	Greg	Scientist	Los Alamos National Laboratory
Scheel	Mark	Professor	California Institute of Technology
Schnetter	Erik	Professor	Perimeter Institute for Theoretical Physics
Siedel	Ed	Professor	University of Illinois Urbana-Champaign
Siegel	Daniel	Professor	Perimeter Institute for Theoretical Physics
Silva	Raniere	Professor	Universtiy of Manchester
Sprouse	Trevor	Postdoctoral Associate	Los Alamos National Laboratory
Surman	Rebecca	Professor	University of Notre Dame
Teukolsky	Saul	Professor	Cornell University
Torres	Christopher	Undergraduate Research Associate	Los Alamos National Laboratory
Vestrand	Thomas	Scientist	Los Alamos National Laboratory
Vincent	Trevor	Unknown	Unknown
Wiggins	Brandon	Professor	University of Southern Utah
Wolfe	Noah	Undergraduate Research Associate	North Carolina State university
Wollaeger	Ryan	Scientist	Los Alamos National Laboratory
Wozniak	Przemyslaw	Scientist	Los Alamos National Laboratory

Identification of Potential Conflicts of Interest or Bias in Selection of Reviewers

N/A

Appendix 2: Current and Pending Support

NSF CURRENT AND PENDING SUPPORT

PI/co-PI/Senior Personnel: Shen, Han-Wei

PROJECT/PROPOSAL CURRENT SUPPORT

1. Project/Proposal Title: RAPIDS2: A SciDAC institute for computer science, data, and artificial intelligence

   Proposal/Award Number (if available): DE-SC0021360

   Source of Support: US DOE

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2020/09

   Project/Proposal Support End Date (if available): 2025/09

   Total Award Amount (including Indirect Costs): $550,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2021	1
2022	1
2023	1
2024	0.92

2. Project/Proposal Title: III: Medium: Collaborative Research: Deep learning for in situ analysis and visualization

   Proposal/Award Number (if available): 1955764

   Source of Support: NSF

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2020/06

   Project/Proposal Support End Date (if available): 2024/05

   Total Award Amount (including Indirect Costs): $715,314

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2020	0.5
2021	1
2022	1
2023	1
2024	0.5

3. Project/Proposal Title: A SciDAC institute for computer science and data

   Proposal/Award Number (if available): 4000159557

   Source of Support: UT-Battelle LLC (Prime: US DOE)

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2018/02

   Project/Proposal Support End Date (if available): 2021/09

   Total Award Amount (including Indirect Costs): $299,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2020	0.5
2021	0.5

4. Project/Proposal Title: Visual analytics for large scale scientific ensemble datasets

   Proposal/Award Number (if available): 471415

   Source of Support: Los Alamos Nat Lab

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2018/02

   Project/Proposal Support End Date (if available): 2021/09

   Total Award Amount (including Indirect Costs): $600,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed

Year	Person-months per year committed
2020	0.5
2021	0.5

5. Project/Proposal Title: RIDIR: Survey data recycling: New analytic framework, integrated database, and tools for cross-national social, behavioral and economic research

   Proposal/Award Number (if available): 1738502

   Source of Support: NSF

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2017/09

   Project/Proposal Support End Date (if available): 2022/08

   Total Award Amount (including Indirect Costs): $1,402,259

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2020	0.27
2021	0.27

PROJECT/PROPOSAL PENDING SUPPORT

1. Project/Proposal Title: Visual analytics of query-based document retrieval to support information seeking for evidence-based practice

   Proposal/Award Number (if available):

   Source of Support: Washington University (Prime: NIH)

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2022/04

   Project/Proposal Support End Date (if available): 2027/03

   Total Award Amount (including Indirect Costs): $998,419

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed

Year	Person-months per year committed
2023	1.35
2024	1.35
2025	1.35
2026	1.35
2027	1.35

2. Project/Proposal Title: Multiscale Data Reduction for Analysis and Visualization of Extreme-Sca Particle Data via Scientific Machine Learning

   Proposal/Award Number (if available):

   Source of Support: DOE

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2021/09

   Project/Proposal Support End Date (if available): 2024/08

   Total Award Amount (including Indirect Costs): $607,665

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2022	1.25
2023	1.25
2024	1.25

3. Project/Proposal Title: AI Institute: ICICLE: Intelligent CyberInfrastructure with Computational Learning in the Environment

   Proposal/Award Number (if available):

   Source of Support: NSF

   Primary Place of Performance: OSU

   Project/Proposal Support Start Date (if available): 2021/07

   Project/Proposal Support End Date (if available): 2026/06

   Total Award Amount (including Indirect Costs): $19,999,997

Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2022	0.5
2023	0.5
2024	0.5
2025	0.5
2026	0.5

*PI/co-PI/Senior Personnel Name: Chris Fryer

***Required fields**

Note: NSF has provided 15 project/proposal and 10 in-kind contribution entries for users to populate. Please leave any unused entries blank.

Project/Proposal Section:

Current and Pending Support includes all resources made available to an individual in support of and/or related to all of his/her research efforts, regardless of whether or not they have monetary value.[1] Information must be provided about all current and pending support, including this project, for ongoing projects, and for any proposals currently under consideration from whatever source[2], irrespective of whether such support is provided through the proposing organization or is provided directly to the individual. Concurrent submission of a proposal to other organizations will not prejudice its review by NSF, if disclosed.[3]

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current, Pending, Submission Planned, and Transfer of Support from top to bottom

[1] If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.
[2] For example, Federal, State, local, foreign, public or private foundations, non-profits, industrial or other commercial organizations or internal funds allocated toward specific projects.
[3] The Biological Sciences Directorate exception to this policy is delineated in PAPPG Chapter II.D.2.

Projects/Proposals

1. *Project/Proposal Title : TEAMS project

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available):

*Source of Support: DOE SciDAC

*Primary Place of Performance : PI is at ORNL, Large collaboration, I only list LANL funding

Project/Proposal Start Date (MM/YYYY) (if available) : 09/2017

Project/Proposal End Date (MM/YYYY) (if available) : 09/2022

*Total Award Amount (including Indirect Costs): $ 500,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2018	0.00
2. 2019	0.00
3. 2020	0.00

Year (YYYY)	Person Months (##.##)
4. 2021	0.00
5. 2022	0.00

2. *Project/Proposal Title : WoU-MMA Collaborative Research: Constraining the nuclear EOS and neutron star astrophysics through multi messenger and multi-object observations of neutron stars

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available): Proposal 1909534

*Source of Support: NSF

*Primary Place of Performance : Rochester Institute of Technology, no funding to LANL

Project/Proposal Start Date (MM/YYYY) (if available) : 09/2019

Project/Proposal End Date (MM/YYYY) (if available) : 09/2022

*Total Award Amount (including Indirect Costs): $ 0

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2019	0.00
2. 2020	0.00
3. 2021	0.00

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

3. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

4. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

5. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

6. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

7. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

8. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

9. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

10. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

11. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

12. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

13.*Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

14.*Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

15. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

In Kind Contributions

*Required fields

In-Kind Contribution Section:

Current and Pending Support also includes in-kind contributions (such as office/laboratory space, equipment, supplies, employees, students). If the in-kind contributions are intended for use on the project being proposed to NSF, the information must be included as part of the Facilities, Equipment and Other Resources section of the proposal and need not be replicated in the individual's Current and Pending Support submission. In-kind contributions not intended for use on the project/proposal being proposed that have associated time obligations must be reported below. If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current to Pending from top to bottom

1.*Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

2. *Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

3. *Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

4.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1.		4.	
2.		5.	
3.			

*Dollar Value of In-Kind Contribution: $

5.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1.		4.	
2.		5.	
3.			

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

6. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

7. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

8.*Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

9.*Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

10. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

*PI/co-PI/Senior Personnel Name: Ayan Biswas

***Required fields**

Note: NSF has provided 15 project/proposal and 10 in-kind contribution entries for users to populate. Please leave any unused entries blank.

Project/Proposal Section:

Current and Pending Support includes all resources made available to an individual in support of and/or related to all of his/her research efforts, regardless of whether or not they have monetary value.[1] Information must be provided about all current and pending support, including this project, for ongoing projects, and for any proposals currently under consideration from whatever source[2], irrespective of whether such support is provided through the proposing organization or is provided directly to the individual. Concurrent submission of a proposal to other organizations will not prejudice its review by NSF, if disclosed.[3]

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current, Pending, Submission Planned, and Transfer of Support from top to bottom

[1] If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.
[2] For example, Federal, State, local, foreign, public or private foundations, non-profits, industrial or other commercial organizations or internal funds allocated toward specific projects.
[3] The Biological Sciences Directorate exception to this policy is delineated in PAPPG Chapter II.D.2.

Projects/Proposals

1. *Project/Proposal Title : ALPINE (Algorithms and Infrastructure for In Situ Visualization and Analysis/ZFP) under ECP (Exascale Computing Project)

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available):

*Source of Support: DOE

*Primary Place of Performance : Los Alamos National Laboratory

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $ 844,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2021	5
2. 2022	4
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

2. *Project/Proposal Title : In-Situ Inference: Bringing Advanced Data Science Into Exascale Simulations

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available):

*Source of Support: LANL LDRD

*Primary Place of Performance : Los Alamos National Laboratory

Project/Proposal Start Date (MM/YYYY) (if available) : 10/2019

Project/Proposal End Date (MM/YYYY) (if available) : 09/2022

*Total Award Amount (including Indirect Costs): $ 1,600,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2021	5
2. 2022	5
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

3. *Project/Proposal Title : Identification of Material Model Parameters using Advanced Learning

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available):

*Source of Support: LANL ASC/PEM

*Primary Place of Performance : Los Alamos National Laboratory

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $ 400,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2021	2
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

4. *Project/Proposal Title : Exascale Spatio-temporal Bayesian Inference

*Status of Support : ( ) Current (•) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available):

*Source of Support: Office of Science - ASCR - DE-FOA-0002493

*Primary Place of Performance : Los Alamos National Laboratory

Project/Proposal Start Date (MM/YYYY) (if available) : 10/2021

Project/Proposal End Date (MM/YYYY) (if available) : 09/2024

*Total Award Amount (including Indirect Costs): $ 380,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2022	3
2. 2023	3
3. 2024	3

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

5. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

6. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

7. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

8. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

9. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

10. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

11. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

12. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

13. *Project/Proposal Title :

*Status of Support : ◯ Current ◯ Pending ◯ Submission Planned ◯ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

14. *Project/Proposal Title :

*Status of Support : ◯ Current ◯ Pending ◯ Submission Planned ◯ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

15. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

In Kind Contributions

*Required fields

In-Kind Contribution Section:

Current and Pending Support also includes in-kind contributions (such as office/laboratory space, equipment, supplies, employees, students). If the in-kind contributions are intended for use on the project being proposed to NSF, the information must be included as part of the Facilities, Equipment and Other Resources section of the proposal and need not be replicated in the individual's Current and Pending Support submission. In-kind contributions not intended for use on the project/proposal being proposed that have associated time obligations must be reported below. If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current to Pending from top to bottom

1.*Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

2.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

3.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

4.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

5.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

6. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

7. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

8. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

9. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

10. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

*PI/co-PI/Senior Personnel Name: Jonah Miller

***Required fields**

Note: NSF has provided 15 project/proposal and 10 in-kind contribution entries for users to populate. Please leave any unused entries blank.

Project/Proposal Section:

Current and Pending Support includes all resources made available to an individual in support of and/or related to all of his/her research efforts, regardless of whether or not they have monetary value.[1] Information must be provided about all current and pending support, including this project, for ongoing projects, and for any proposals currently under consideration from whatever source[2], irrespective of whether such support is provided through the proposing organization or is provided directly to the individual. Concurrent submission of a proposal to other organizations will not prejudice its review by NSF, if disclosed.[3]

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current, Pending, Submission Planned, and Transfer of Support from top to bottom

[1] If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.
[2] For example, Federal, State, local, foreign, public or private foundations, non-profits, industrial or other commercial organizations or internal funds allocated toward specific projects.
[3] The Biological Sciences Directorate exception to this policy is delineated in PAPPG Chapter II.D.2.

Projects/Proposals

1.*Project/Proposal Title : TEAMS project

*Status of Support : ◉ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support: DOE SciDAC

*Primary Place of Performance : PI is at ORNL, Large collaboration, I only list LANL funding

Project/Proposal Start Date (MM/YYYY) (if available) : 09/2017

Project/Proposal End Date (MM/YYYY) (if available) : 09/2022

*Total Award Amount (including Indirect Costs): $ 500,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1. 2017	6.00	4. 2021	0.00
2. 2018	3.00	5. 2022	0.00
3. 2019	0.00		

2.*Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1.		4.	
2.		5.	
3.			

Projects/Projects/Proposals

3. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

4. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

5. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

6. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

7. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

8. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

9. *Project/Proposal Title :

*Status of Support : ◯ Current ◯ Pending ◯ Submission Planned ◯ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

10. *Project/Proposal Title :

*Status of Support : ◯ Current ◯ Pending ◯ Submission Planned ◯ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

11. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

12. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

13. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

14. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

Projects/Proposals

15. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Required fields

In-Kind Contribution Section:

Current and Pending Support also includes in-kind contributions (such as office/laboratory space, equipment, supplies, employees, students). If the in-kind contributions are intended for use on the project being proposed to NSF, the information must be included as part of the Facilities, Equipment and Other Resources section of the proposal and need not be replicated in the individual's Current and Pending Support submission. In-kind contributions not intended for use on the project/proposal being proposed that have associated time obligations must be reported below. If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current to Pending from top to bottom

1. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

2. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

3. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

4. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

5. *Status of Support : ○ Current ○ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

6.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

7.*Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

8. *Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

9. *Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

In Kind Contributions

10. *Status of Support : ◯ Current ◯ Pending

*Source of Support :

*Primary Place of Performance :

*Summary of In-Kind Contributions :

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

*Year (YYYY)	*Person Months (##.##)
1.	
2.	
3.	

Year (YYYY)	Person Months (##.##)
4.	
5.	

*Dollar Value of In-Kind Contribution: $

Appendix 3: Bibliography and References Cited

References

[1] SO/IEC 15444-1. Information technology–JPEG 2000 image coding system, 2000. Standard, International Organization for Standardization.

[2] B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. , 848(2):L12, October 2017.

[3] Yifei An, Han-Wei Shen, and Guan Li. STSRNet: Deep joint space-time super-resolutionfor vector field visualization. *Submitted for publication*.

[4] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein Generative Adversarial Networks. In *Proc. 2017 International Conference on Machine Learning*, pages 214–223, 2017.

[5] Johannes Ballé, Valero Laparra, and Eero P. Simoncelli. End–to–end optimized image compression. In *Proc. 2017 International Conference on Learning Representations*, 2017.

[6] Wenbo Bao, Wei-Sheng Lai, Xiaoyun Zhang, Zhiyong Gao, and Ming-Hsuan Yang. Memc-net: Motion estimation and motion compensation driven neural network for video interpolation and enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, page 1–1, 2019.

[7] L. Becerra, C. L. Ellinger, C. L. Fryer, J. A. Rueda, and R. Ruffini. SPH Simulations of the Induced Gravitational Collapse Scenario of Long Gamma-Ray Bursts Associated with Supernovae. , 871(1):14, January 2019.

[8] Matthew Berger, Jixian Li, and Joshua Aaron Levine. A generative model for volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 25(4):1636–1650, 2019.

[9] R. D. Blandford and R. L. Znajek. Electromagnetic extraction of energy from Kerr black holes. *Monthly Notices of the Royal Astronomical Society*, 179(3):433–456, 07 1977.

[10] Udeepta Bordoloi and Han-Wei Shen. View selection for volume rendering. In *16th IEEE Visualization Conference, VIS 2005, Minneapolis, MN, USA, October 23-28, 2005*, pages 487–494. IEEE Computer Society, 2005.

[11] G. E. P. Box and Mervin E. Muller. A Note on the Generation of Random Normal Deviates. *The Annals of Mathematical Statistics*, 29(2):610 – 611, 1958.

[12] Zhutian Chen, Wei Zeng, Zhiguang Yang, Lingyun Yu, Chi Wing Fu, and Huamin Qu. LassoNet: Deep Lasso-Selection of 3D Point Clouds. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):195–204, 2020.

[13] Hsueh-Chien Cheng, Antonio Cardone, Somay Jain, Eric Krokos, Kedar Narayan, Sriram Subramaniam, and Amitabh Varshney. Deep-learning-assisted volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 25(2):1378–1391, 2019.

[14] Theodoros Christoudias, Christos Kallidonis, Loizos Koutsantonis, Christos Lemesios, Lefteris Markou, and Constantinos Sophocleous. Visualising the dark sky IEEE SciVis contest 2015. *2015 IEEE Scientific Visualization Conference, SciVis 2015 - Proceedings*, pages 79–86, 2016.

[15] B. Côté et al. The Origin of r-process Elements in the Milky Way. , 855:99, March 2018.

[16] Arthur Dempster, Natalie Laird, and D.B. Rubin. Maximum likelihood from incomplete data via em algorithm. *J. Royal Statistical Soc., Series B*, 39:1 – 38, 09 1977.

[17] Soumya Dutta, Chun-Ming Chen, Gregory Heinlein, Han-Wei Shen, and Jen-Ping Chen. In situ distribution guided analysis and visualization of transonic jet engine simulations. *IEEE transactions on visualization and computer graphics*, 23(1):811–820, 2017.

[18] Soumya Dutta, Xiaotong Liu, Ayan Biswas, Han-Wei Shen, and Jen-Ping Chen. Pointwise information guided visual analysis of time-varying multi-fields. In Koji Koyamada and Puripant Ruchikachorn, editors, *SIGGRAPH ASIA 2017, Bangkok, Thailand, November 27 - 30, 2017 - Symposium on Visualization*, pages 17:1–17:8. ACM, 2017.

[19] David Eichler, Mario Livio, Tsvi Piran, and David N. Schramm. *Nucleosynthesis, neutrino bursts and γ-rays from coalescing neutron stars*, pages 682–684. 1996.

[20] Francois Foucart. Monte Carlo closure for moment-based transport schemes in general relativistic radiation hydrodynamic simulations. , 475(3):4186–4207, April 2018.

[21] Francois Foucart, Matthew D. Duez, Francois Hebert, Lawrence E. Kidder, Harald P. Pfeiffer, and Mark A. Scheel. Monte-carlo neutrino transport in neutron star merger simulations. *The Astrophysical Journal*, 902(1):L27, oct 2020.

[22] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Super–Resolution Reconstruction of Turbulent Flows With Machine Learning. *Journal of Fluid Mechanics*, 870:106–120, 2019.

[23] Kai Fukami, Koji Fukagata, and Kunihiko Taira. Machine–learning-based spatio–temporal super resolution reconstruction of turbulent flows. *Journal of Fluid Mechanics*, 909:A9, 2021.

[24] Patrick Gralka, Sebastian Grottel, Joachim Staib, Karsten Schatz, Grzegorz Karch, Manuel Hirschler, Michael Krone, Guido Reina, Stefan Gumhold, and Thomas Ertl. 2016 IEEE Scientific Visualization Contest Winner: Visual and Structural Analysis of Point-based Simulation Ensembles. *IEEE Computer Graphics and Applications*, 38(3):106–117, 2018.

[25] L. Guo, S. Ye, J. Han, H. Zheng, H. Gao, D. Z. Chen, J. Wang, and C. Wang. SSR–VFD: Spatial Super-Resolution for Vector Field Data Analysis and Visualization. In *Proc. 2020 IEEE Pacific Visualization Symposium*, pages 71–80, 2020.

[26] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennamoun. Deep Learning for 3D Point Clouds: A Survey. 8828(c):1–27, 2019.

[27] J. Han and C. Wang. SSR–TVD: Spatial Super-Resolution for Time-Varying Data Analysis and Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2020.

[28] J. Han and C. Wang. TSR–TVD: Temporal Super–Resolution for Time-Varying Data Analysis and Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):205–215, 2020.

[29] Fiona A. Harrison et al. The Nuclear Spectroscopic Telescope Array (NuSTAR) High-energy X-Ray Mission. , 770(2):103, June 2013.

[30] Kaveh Hassani and Mike Haley. Unsupervised multi-task feature learning on point clouds. *Proceedings of the IEEE International Conference on Computer Vision*, 2019-Octob:8159–8170, 2019.

[31] Wenbin He, Junpeng Wang, Hanqi Guo, Ko-Chih Wang, Han-Wei Shen, Mukund Raj, Youssef S. G. Nashed, and Tom Peterka. InSituNet: Deep image synthesis for parameter space exploration of ensemble simulations. *IEEE Trans. Comput. Graph.*, 26(1):23–33, 2020.

[32] Pedro Hermosilla, Tobias Ritschel, Pere Pau Vázquez, Àlvar Vinacua, and Timo Ropinski. Monte Carlo convolution for learning on non-uniformly sampled point clouds. *arXiv*, 37(6), 2018.

[33] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[34] F. Hong, C. Liu, and X. Yuan. DNN-VolVis: Interactive volume visualization supported by deep neural network. In *Proceedings of IEEE Pacific Visualization Symposium*, 2019.

[35] Jakob Jakob, Markus Gross, and Tobias Günther. A Fluid Flow Data Set for Machine Learning and its Application to Neural Flow Map Interpolation. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1279–1289, 2021.

[36] C. Jiang, S. Esmaeilzadeh, K. Azizzadenesheli, K. Kashinath, M. Mustafa, H. Tchelepi, P. Marcus, M. Prabhat, and A. Anandkumar. MESHFREEFLOWNET: A Physics-Constrained Deep Continuous Space–Time Super-Resolution Framework. In *Proc. 2020 International Conference for High Performance Computing, Networking, Storage and Analysis*, number 9, pages 1–15. IEEE Computer Society, 2020.

[37] Huaizu Jiang, Deqing Sun, Varan Jampani, Ming-Hsuan Yang, Erik Learned-Miller, and Jan Kautz. Super slomo: High quality estimation of multiple intermediate frames for video interpolation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun 2018.

[38] J. A. Kegerreis, L. F. A. Teodoro, V. R. Eke, R. J. Massey, D. C. Catling, C. L. Fryer, D. G. Korycansky, M. S. Warren, and K. J. Zahnle. Consequences of Giant Impacts on Early Uranus for Rotation, Internal Structure, Debris, and Atmospheric Erosion. , 861(1):52, July 2018.

[39] C. A. Kierans, S. E. Boggs, A. Zoglauer, A. W. Lowell, C. Sleator, J. Beechert, T. J. Brandt, P. Jean, H. Lazar, J. Roberts, T. Siegert, J. A. Tomsick, and P. von Ballmoos. Detection of the 511 keV Galactic Positron Annihilation Line with COSI. , 895(1):44, May 2020.

[40] Shiyi Lan, Ruichi Yu, Gang Yu, and Larry S. Davis. Modeling local geometric structure of 3D point clouds using geo-CNN. *arXiv*, pages 998–1008, 2018.

[41] Haoyu Li and Han-Wei Shen. Time-varying particle feature extraction and tracking using neural networks. *Submitted for publication*.

[42] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. PointCNN: Convolution on X-transformed points. *Advances in Neural Information Processing Systems*, 2018-Decem:820–830, 2018.

[43] J. Lippuner and L. F. Roberts. SkyNet: A Modular Nuclear Reaction Network Library. , 233:18, December 2017.

[44] Shusen Liu, Joshua A Levine, Peer-Timo Bremer, and Valerio Pascucci. Gaussian mixture model based volume visualization. In *IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, pages 73–77. IEEE, 2012.

[45] Xiaotong Liu and Han-Wei Shen. Association analysis for visual exploration of multivariate scientific data sets. *IEEE Trans. Vis. Comput. Graph.*, 22(1):955–964, 2016.

[46] Kwan-Liu Ma. Machine learning to boost the next generation of visualization technology. *IEEE Computer Graphics and Applications*, 27(5):6–9, 2007.

[47] Donald Meagher. Geometric Modeling Using Octree Encoding. *Computer Graphics and Image Processing*, 19(2):129–147, 1982.

[48] F. Melia. *High-energy Astrophysics*. High-energy Astrophysics. Princeton University Press, 2009.

[49] Jonah M. Miller, Ben. R. Ryan, and Joshua C. Dolence. ν bhlight: Radiation GRMHD for Neutrino-driven Accretion Flows. , 241(2):30, April 2019.

[50] Jonah M. Miller, Benjamin R. Ryan, Joshua C. Dolence, Adam Burrows, Christopher J. Fontes, Christopher L. Fryer, Oleg Korobkin, Jonas Lippuner, Matthew R. Mumpower, and Ryan T. Wollaeger. Full transport model of gw170817-like disk produces a blue kilonova. *Phys. Rev. D*, 100:023008, Jul 2019.

[51] Jonah M. Miller, Trevor M. Sprouse, Christopher L. Fryer, Benjamin R. Ryan, Joshua C. Dolence, Matthew R. Mumpower, and Rebecca Surman. Full transport general relativistic radiation magnetohydrodynamics for nucleosynthesis in collapsars. *The Astrophysical Journal*, 902(1):66, oct 2020.

[52] J. J. Monaghan. Smoothed particle hydrodynamics. , 30:543–574, January 1992.

[53] Patrick M. Motl, Juhan Frank, Jan Staff, Geoffrey C. Clayton, Christopher L. Fryer, Wesley Even, Steven Diehl, and Joel E. Tohline. A Comparison of Grid-based and SPH Binary Mass-transfer and Merger Simulations. , 229(2):27, April 2017.

[54] M. R. Mumpower, T. Kawano, J. L. Ullmann, M. Krtička, and T. M. Sprouse. Estimation of M 1 scissors mode strength for deformed nuclei in the medium- to heavy-mass region by statistical Hauser-Feshbach model calculations. *Physical Review C*, 96(2), August 2017.

[55] Jean-Claude Passy, Orsola De Marco, Chris L. Fryer, Falk Herwig, Steven Diehl, Jeffrey S. Oishi, Mordecai-Mark Mac Low, Greg L. Bryan, and Gabriel Rockefeller. Simulating the Common Envelope Phase of a Red Giant Using Smoothed-particle Hydrodynamics and Uniform-grid Codes. , 744(1):52, January 2012.

[56] David Posada and Thomas R Buckley. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic biology*, 53(5):793–808, 2004.

[57] Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. PointNet: Deep learning on point sets for 3D classification and segmentation. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:77–85, 2017.

[58] Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems*, 2017-Decem:5100–5109, 2017.

[59] Mohammad Raji, Alok Hota, Robert Sisneros, Peter Messmer, and Jian Huang. Photo-guided exploration of volume data features. In *Proceedings of Eurographics Symposium on Parallel Graphics and Visualization*, pages 31–39, 2017.

[60] Tobias Rapp, Christoph Peters, and Carsten Dachsbacher. Visual Analysis of Large Multivariate Scattered Data using Clustering and Probabilistic Summaries. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2020.

[61] Sherwood Richers, Daniel Kasen, Evan O'Connor, Rodrigo Fernández, and Christian D. Ott. Monte Carlo Neutrino Transport through Remnant Disks from Neutron Star Mergers. , 813(1):38, November 2015.

[62] Gabriel Rockefeller, Christopher L. Fryer, and Fulvio Melia. Spin-induced Disk Precession in Sagittarius A*. , 635(1):336–340, December 2005.

[63] Gabriel Rockefeller, Christopher L. Fryer, Fulvio Melia, and Q. Daniel Wang. Diffuse X-Rays from the Arches and Quintuplet Clusters. , 623(1):171–180, April 2005.

[64] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, page 234–241, 2015.

[65] Murray Rosenblatt. Remarks on some non-parametric estimates of a density function. *The Annals of Mathematical Statistics*, 27, 09 1956.

[66] Mateu Sbert, Han-Wei Shen, Ivan Viola, Min Chen, Anton Bardera, and Miquel Feixas. Tutorial on information theory in visualization. In Tomasz Bednarz, editor, *SIGGRAPH Asia 2017 Courses, Bangkok, Thailand, November 27 - 30, 2017*, pages 17:1–17:165. ACM, 2017.

[67] Karsten Schatz, Christoph Müller, Patrick Gralka, Moritz Heinemann, Alexander Straub, Christoph Schulz, Matthias Braun, Tobias Rau, Michael Becher, Steffen Frey, et al. 2019 ieee scientific visualization contest winner: Visual analysis of structure formation in cosmic evolution. *IEEE Computer Graphics and Applications*, 2020.

[68] Neng Shi and Yubo Tao. CNNs based viewpoint estimation for volume visualization. *ACM Transactions on Intelligent Systems and Technology*, 10(3):27:1–27:22, 2019.

[69] T. M. Sprouse, R. Navarro Perez, R. Surman, M. R. Mumpower, G. C. McLaughlin, and N. Schunck. Propagation of Statistical Uncertainties of Skyrme Mass Models to Simulations of r-Process Nucleosynthesis. *submitted to Physical Review*, March 2019. arXiv: 1901.10337.

[70] Jan. E. Staff, Brandon Wiggins, Dominic Marcello, Patrick M. Motl, Wesley Even, Chris L. Fryer, Cody Raskin, Geoffrey C. Clayton, and Juhan Frank. The Role of Dredge-up in Double White Dwarf Mergers. , 862(1):74, July 2018.

[71] W. Tao, F. Jiang, S. Zhang, J. Ren, W. Shi, W. Zuo, X. Guo, and D. Zhao. An End–to–End Compression Framework Based on Convolutional Neural Networks. In *Proc. 2017 Data Compression Conference*, pages 463–463, 2017.

[72] L. Theis, W. Shi, A. Cunningham, and F. Huszár. Lossy image compression with compressive autoencoders. In *Proc. 2017 International Conference on Learning Representations*, 2017.

[73] Hugues Thomas, Charles R. Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, François Goulette, and Leonidas J. Guibas. KPConv: Flexible and Deformable Convolution for Point Clouds. 2019.

[74] G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell. Full Resolution Image Compression with Recurrent Neural Networks. In *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pages 5435–5443, 2017.

[75] George Toderici, Sean M. O'Malley, Sung Jin Hwang, Damien Vincent, David Minnen, Shumeet Baluja, Michele Covell, and Rahul Sukthankar. Variable Rate Image Compression with Recurrent Neural Networks. In *Proc. 2016 International Conference on Learning Representations*, 2016.

[76] Fan-Yin Tzeng, Eric B. Lum, and Kwan-Liu Ma. A novel interface for higher-dimensional classification of volume data. In *Proc. IEEE Visualization '03*, pages 505–512, 2003.

[77] Fan-Yin Tzeng, Eric B. Lum, and Kwan-Liu Ma. An intelligent system approach to higher-dimensional classification of volume data. *IEEE Trans. Vis. Comput. Graph.*, 11(3):273–284, 2005.

[78] Chaoli Wang and Han-Wei Shen. LOD map - A visual interface for navigating multiresolution volume visualization. *IEEE Trans. Vis. Comput. Graph.*, 12(5):1029–1036, 2006.

[79] Chaoli Wang and Han-Wei Shen. Information theory in scientific visualization. *Entropy*, 13(1):254–273, 2011.

[80] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic graph Cnn for learning on point clouds. *ACM Transactions on Graphics*, 38(5), 2019.

[81] Z. Wang, J. Chen, and S. C. H. Hoi. Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, to appear.

[82] Tzu-Hsuan Wei, Soumya Dutta, and Han-Wei Shen. Information guided data sampling and recovery using bitmap indexing. In *IEEE Pacific Visualization Symposium, PacificVis 2018, Kobe, Japan, April 10-13, 2018*, pages 56–65. IEEE Computer Society, 2018.

[83] Tzu-Hsuan Wei, Teng-Yok Lee, and Han-Wei Shen. Evaluating isosurfaces with level-set-based information maps. *Comput. Graph. Forum*, 32(3):1–10, 2013.

[84] S. Weiss, M. Chu, N. Thuerey, and R. Westermann. Volumetric isosurface rendering with deep learning-based super-resolution. *arXiv preprint arXiv:1906.06520*, 2019.

[85] Wenxuan Wu, Zhongang Qi, and Li Fuxin. PointCONV: Deep convolutional networks on 3D point clouds. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-June:9613–9622, 2019.

[86] Skylar Wurster, Han-Wei Shen, Hanqi Guo, and Thomas Peterka. Deep hierarchical super-resolution for scientific data reduction and visualization. *Submitted for publication*.

[87] You Xie, Erik Franz, Mengyu Chu, and Nils Thuerey. tempoGAN: A Temporally Coherent, Volumetric GAN for Super–Resolution Fluid Flow. *ACM Transactions on Graphics*, 37(4):95:1–95:15, 2018.

[88] Lijie Xu, Teng-Yok Lee, and Han-Wei Shen. An information-theoretic framework for flow visualization. *IEEE Trans. Vis. Comput. Graph.*, 16(6):1216–1224, 2010.

[89] Zhenglei Zhou, Yule Hou, Qirui Wang, Guangxiang Chen, Jiawei Lu, Yubo Tao, and Hai Lin. Volume upscaling with convolutional neural networks. In *Proc. 2017 Computer Graphics International Conference*, pages 1–6, 2017.

[90] Zhenglei Zhou, Yule Hou, Qirui Wang, Guangxiang Chen, Jiawei Lu, Yubo Tao, and Hai Lin. Volume upscaling with convolutional neural networks. In *Proceedings of Computer Graphics International*, pages 38:1–38:6, 2017.

[91] Y. Zhu, R. T. Wollaeger, N. Vassh, R. Surman, T. M. Sprouse, M. R. Mumpower, P. Möller, G. C. McLaughlin, O. Korobkin, T. Kawano, P. J. Jaffke, E. M. Holmbeck, C. L. Fryer, W. P. Even, A. J. Couture, and J. Barnes. Californium-254 and Kilonova Light Curves. *The Astrophysical Journal*, 863(2):L23, August 2018.

Appendix 4: Facilities and Other Resources

The Ohio State University

The critical requirement will be high performance computing resources to run fine-grained simulations and per- form analyses, including on accelerators like GPUs, and for certain studies, visualization devices. The PI has access to resources in the department of Computer Science and Engineering (CSE), Mechanical and Aerospace Engineering (MEA), the Ohio Supercomputing Center (OSC), as well as major NSF and DOE resources.

The relevant resources at OSC include: (1) the Oakley HP Intel Xeon cluster, which features total 8,328 cores on 694 nodes with 12 cores and 48 GB DRAM per node, and (2) Glenn IBM Cluster 1350, which consists of 658 System x3455 compute nodes with dual socket, quad core 2.5 GHz Opteron and 24 GB DRAM per node. Many of the nodes at both the clusters have different types of GPUs, and plans are to purchase MIC nodes soon.

We can also leverage two research clusters provided through NSF Research Infrastructure (RI) program. One cluster resides in the Department of Computer Science and Engineering. It consists of 64 nodes with dual Intel EM64T processors on the 32 nodes and dual AMD Opteron processors on the other 32 nodes. The nodes are connected through InfiniBand DDR. The other cluster resides in the Department of Biomedical Informatics. It consists of 64 nodes with AMD dual Opteron processors, which are connected via InfiniBand SDR. 8 nodes in the first cluster and 32 nodes in the second cluster have GPUs. Another relevant resource within CSE is in Shen's visualization research lab, which, besides workstations and the GPU, includes a 4x4 display panel wall, where each panel is a 40-inch flat LCD screen. Through PI's projects with DOE and NSF, students either already have or will likely be able to access resources from XSEDE, as well as the supercomputers at Argonne National Laboratory, Oak Ridge National Laboratory, and Lawrence Berkeley National Laboratory.

Los Alamos National Laboratory

LANL provides facilities for external collaborative research projects, on what is known as the *Turquoise* Institutional Computing Network (ICN). *Darwin* is another LANL research computing cluster, available for internal-network and external access through VPN.

Turquoise ICN LANL Institutional Computing provides access to several different computing systems that are tailored to specific classes of workloads. The current offerings include: *Grizzly*, for large-scale workloads, *Badger*, for small- to medium-scale workloads, *Kodiak*, for GPGPU-enabled and machine-learning workloads, and *Capulin*, an experimental ARM system with high memory-bandwidth. Table 4 is a current summary of the machines in the LANL Turquoise ICN. All of the clusters utilize a general-purpose commodity processor (x86-64) and a commercial interconnect. The systems share a single user-space file system, with additional Lustre scratch space resources available per cluster. LANL is also about to undergo a hardware refresh, and we expect new hardware to be purchased for the Turquoise network within the timescale of the grant.

Darwin Darwin is an NNSA ASC and SC ASCR funded test-bed cluster. Of particular interest, Darwin is configured with a non-standard HPC software stack with heterogeneous hardware nodes and software configurations. As such, Darwin does not have uniform and predictable run-time capabilities. Rather, it has a wide variety of hardware and software to test various different hardware configurations, algorithms, and software designs. Table 5 provides a small sampling of the hardware capabilities of various nodes within Darwin. Additionally, the Darwin support-staff is available to install new hardware or software reconfiguration based on user requests.

Turquoise (Open Collaborative Network) - IC

Cluster Name	Processor	OS	Total Compute Nodes	CPU cores per Node / Total CPU cores	Memory per compute Node / Total Memory	Interconnect	Peak (TFlop/s)	Storage
Badger CTS-1	Intel Xeon Broadwell	Linux (TOSS)	660 nodes	36 / 23,760	128 GB / 84.5 TiB	Intel OmniPath	798	23.5 PB Lustre
Capulin ARM	Cavium ThunderX2	SLES-based CLE	175 nodes	56 / 9,800	256 GB / 44.8 TB	Cray Aries	196	23.5 PB Lustre
Grizzly CTS-1	Intel Xeon Broadwell	Linux (TOSS)	1490 nodes	36 / 53,640	128 GB / 191 TB	Intel OmniPath	1,806	23.5 PB Lustre
Kodiak CTS-1	Intel Xeon E5-2695 v4 + 4x NVidia Tesla P100 GPUs	Linux (TOSS)	133 nodes	36 / 4,788 + GPUs	66x512, 67x256 GiB / ~50 TiB	Mellanox InfiniBand EDR Fat-Tree	160 CPU + GPUs	23.5 PB Lustre

Figure 4: Configuration of the collaborative LANL HPC systems.

Node Name	Partition	RAM	CPU Vendor	CPU Family	CPU Model	Base Clock Rate	Sockets	Cores per Socket	Total Cores	Thre Core
Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...	Sea
cn0	go_vnc	62GB	Intel	sandybridge	E5-2660_0	2.20GHz	2	8	16	2
cn0	general	62GB	Intel	sandybridge	E5-2660_0	2.20GHz	2	8	16	2
cn1	general	251GB	Intel	sandybridge	E5-2660_0	2.20GHz	2	8	16	2
cn2	general	251GB	Intel	sandybridge	E5-2660_0	2.20GHz	2	8	16	2
cn5	general	1511GB	Intel	sandybridge	E5-4650L_0	2.60GHz	4	8	32	2
cn6	general	1511GB	Intel	sandybridge	E5-4650L_0	2.60GHz	4	8	32	2
cn7	general	1511GB	Intel	sandybridge	E5-4650L_0	2.60GHz	4	8	32	2
cn8	general	1511GB	Intel	sandybridge	E5-4650L_0	2.60GHz	4	8	32	2
cn21	go_vnc	125GB	Intel	ivybridge	E5-2650_v2	2.60GHz	2	8	16	1
cn21	general	125GB	Intel	ivybridge	E5-2650_v2	2.60GHz	2	8	16	1
cn30	galton-vm	251GB	Intel	ivybridge	E5-2650_v2	2.60GHz	2	8	16	1
cn30	shared-vm	251GB	Intel	ivybridge	E5-2650_v2	2.60GHz	2	8	16	1
cn31	shared-vm	251GB	Intel	ivybridge	E5-2650_v2	2.60GHz	2	8	16	1

Figure 5: A small selection of nodes available on Darwin.

Appendix 5: Equipment

Relevant equipment is included in the discussion of Facilities and Other Resources (Appendix 4).

Appendix 6: Data Management Plan

We recognize the value of preserving a record of our research and sharing the software and data products that result. We believe that well-crafted data management is essential to advancing the quality and pace of scientific research, and it is an integral part of our planning.

We will use public web-based tools designed primarily for software development to manage both software and data artifacts. High-quality systems for building software, managing dependencies, controlling versions, tracking issues, scripting, plotting, building wikis, project pages, and writing documentation are all readily available today. The use of these tools is essential for guaranteeing reproducible results, provided that they are designed into the project from the outset and not as an afterthought. The automation of many of these tasks not only helps others validate our research, it streamlines and organizes our own work and saves us time as well. Sound data management is truly an integral part of an organized scientific method.

This section outlines our strategy for using modern web tools for curating data and software. Preserved artifacts entail source code, documentation, publications, plotting scripts, utilities, test data sets, metadata and notes, and software build scripts. The content, format, and tools managing these data products throughout their life cycle are described below.

Availability. All of the following will be freely available to the public, except in cases where proprietary, business confidential, or personally identifiable information must be protected.

Data products. All source code will be developed and hosted in publicly available repositories such as github, bitbucket, or gitlab. During development, a simple public-domain copyright statement will accompany the software, and when an official release is deemed appropriate, a Berkeley Source Distribution (BSD) style open source license will be sought in accordance with ASCR and Argonne policy. We will use Cmake to maintain the building and installation of our software and to install any needed dependencies. Documentation of the source code will be generated by using Doxygen and will be hosted alongside the code in a public location. Github, for example, provides public web page hosting for this purpose.

Publications in portable document format (PDF) will also reside alongside the software and documentation. Scripts in R or Python to generate plots in publications and the datasets needed to execute them will be similarly curated. Shell commands to generate plots, including their input arguments, will also be stored in Bash or Python scripts, automating their execution and ensuring their reproducibility. We have found Ipython Notebook to also be extremely useful for managing Python scripts for this purpose. Papers that are in preparation will not be publicly available until they have been cleared by the laboratory and submitted for review or have been accepted for publication or have appeared in print or online in final or preprint form.

Additional data sets used in our research consist of small test datasets for validating proof of concept and larger test data for scalable performance testing. In the first case, we will include such sample data sets in the software repository to the extent that there is sufficient space available. In the second case, we will host data sets using laboratory or computing facility resources such as network-mounted and parallel file systems that are maintained by the individual facilities.

Project metadata consists of meeting notes, design documents, project reports, and research highlights. We will use the wiki feature of the software repository to maintain such documents.

Protection. We will ensure that none of the data contains personally identifiable, proprietary, or confidential information. No copyrighted information will be shared without written consent of the copyright holder. No excerpts of scientific data from our collaborating researchers will be shared without their written consent. In all cases, laboratory review and release processes will be followed to ensure that privacy is protected.