Visualization and Analysis Tools for Exploring Magnetic Confinement Fusion Simulations

Allen Sanderson[1], Scott Kruger[2], Stephane Ethier[3], and Joshua Breslau[3]

[1]SCI Institute, University of Utah, Salt Lake City, UT 84112
[2]Tech-X Corporation, Boulder, CO
[3]Princeton Plasma Physics Lab, Princeton, NJ

E-mail: allen@sci.utah.edu, kruger@txcorp.com, ethier@pppl.gov, jbreslau@pppl.gov

Abstract. In this paper, we highlight two analysis tools for evaluating magnetic confinement fusion simulations. Our goal in developing these tools is to allow fusion scientists to perform both in situ and postprocessing analysis of their data. The first tool is for understanding the magnetic field topology in fusion simulations and utilizes a traditional paradigm for analysis. The same analysis tool is used for both in situ and postprocessing to understand the topology. When it is used in situ, scientists can achieve a data reduction at the expense of interaction that can be achieved when postprocessing. The second tool, which allows for query-based analysis and visualization, follows a less traditional paradigm to assist in the analysis of simulation codes that utilize millions to billions of particles. We combine parallel coordinates plots with accelerated index searches to allow scientists to perform queries of their data. Parallel coordinate plots allow one to identify trends within multivariate data, while accelerated index searches allow one to quickly perform range-based queries on a large number of multivariate entries. At the same time we look beyond data analysis and reduction and look towards the future where such tools can be used for computational steering. Such ability will be important as simulations and experiments come closer to their goal of producing a substantiating burning plasma.

1. Introduction

As scientists move to larger simulations, their ability to fully examine each time step will become increasingly limited at both the machine and human level. Even today scientists routinely write out only a small faction ($< 10\%$) of each time step generated. This practice often results in scientists rerunning their simulations when interesting phenomena occur between time steps. In the future, when utilizing exascale machines, scientists may not have this ability. Hence, they will need to rely more heavily on analysis tools to find features of interest during each time step.

2. Identifying Magnetic Islands in Poincaré Sections

In designing efficient fusion reactors scientists must be able to analyze the magnetic field from numerical simulations and characterize its topology. Iin order to achieve a stable plasma equilibrium, the magnetic field should be in the form of a series of nested flux surfaces, Figure 1a. Because of various instabilities [3], however, the magnetic field forms islands chains that may overlap to produce stochasticity, leading to a loss in thermal energy confinement; see Figure 1b. Our first analysis tool is designed to identify topological features (island chains) within the magnetic field.

In Figure 1, we show two Poincaré maps that offer a principled dimensionality reduction to study Hamiltonian systems [5] such as the magnetic field. The map is formed by the intersection of magnetic fieldlines with a plane perpendicular to the axis of the torus, whereby a sufficient number of intersections (i.e., puncture points) are collected in order to reveal salient features. Traditionally, constructing a Poincaré map can best be described as being brute force (i.e., computationally inefficient). That is, a magnetic fieldline is traced (i.e., integrated) regardless of its location or topology. As such, some features are well defined while others are not. Further, being a point-based approach, it relies on the human eye to form a contiguous representation of the magnetic field's cross-sectional profile.

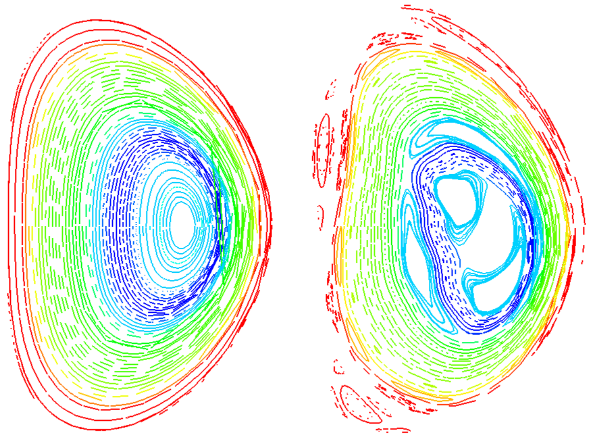


Figure 1. (a) A Poincaré section of the magnetic field of an early time slice from a NIMROD simulation showing nice magnetic confinement of the plasma. (b) A later time slice from the same simulation where the magnetic field has become stochastic, forming islands chains leading to a loss in plasma confinement at the core.

In Refs. [8] and [9], we described a system for addressing these issues by using a technique that analyzes the fieldline's topology while the integration is being performed. We briefly discuss the analysis and refer the reader to these references for further details.

For the moment, consider only quasi-periodic fieldlines that occur for both flux surfaces and island chains. Every fieldline, regardless of the topology, has a toroidal and poloidal periodicity. Their ratio is referred to as the magnetic surface's safety factor, which gives an indication of the stability of the plasma confined by the magnetic field. For a flux-surface, there is an infinite number of pairs of periods (expressed as rational values) that, when expressed as a ratio, approximates the safety factor. For an island chain there is a limited number of periods. Further, we have identified that within these periods there are resonance periods that occur only for island chains. By identifying the resonance periods we can identify island chains and thus the topology. Topologically a primary resonance is associated with a simple island chain, whereas a secondary resonance is associated with satellite islands; see Figure 2.

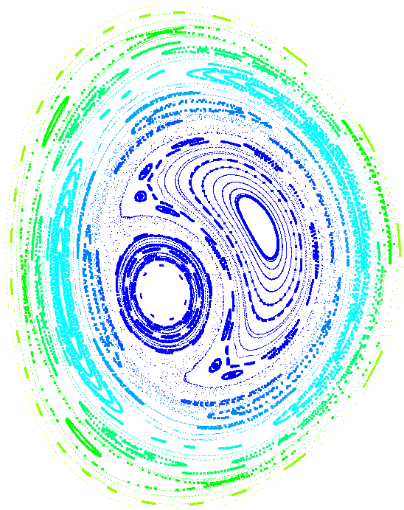
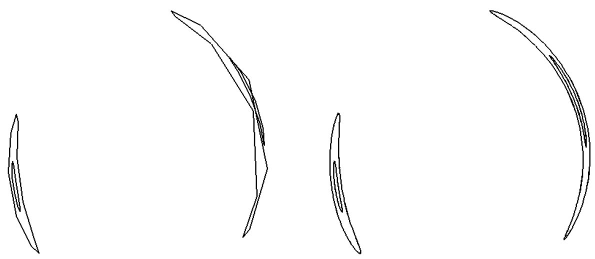


Figure 2. A Poincaré section of the magnetic field of a time slice from a M3D-C1 sawtooth simulation showing a large 1,1 island (primary resonance) with 19 smaller satellite islands (secondary resonance) surrounding it.

Once the topology has been identifieda, we have additional information that can be used to limit the fieldline integration. For an island chain, embedded within the periods is a dominant harmonic period that is an integer multiple of the resonant period. We have observed that this integer multiple is directly related to the number of points that make up a single transit of an island. For example, if the resonant period is 2 and the dominate harmonic is 20, then there will be two islands with each island having 10 points around its boundary; see Figure 3a. Flux surfaces are more difficult, and we have resorted to a geometric process to limit the amount of integration required to fully outline the surface [9]. While this technique allows us to limit the integration on a case-by-case basis and is certainly more advantageous than a brute-force approach, it does have a potential drawback. Instead of relying on a dense set of points for forming a contiguous cross-sectional representation, we connect the points using a series of linear segments. If the distance between the points is small enough, the linear segments give a good approximation to the actual cross section. Unfortunately, this is not always the case, as demonstrated in Figure 3a. While we have some techniques for reducing the size of the segments (Figure 3b), it has not been automated and is part of our future work.

One of the unique features of the analysis is that it is simulation code independent and requires only the magnetic fieldline. This independence is important because many MHD codes simulate the magnetic field, each having its own unique internal element representation. For instance, in Figure 1, we utilized the NIMROD code [10], whereas in Figure 2, we utilized the M3D-C1 code [2].



(a) (b)

Figure 3. (a) A Poincaré section showing a nested 2,1 island chain using a minimal number of segments (10) to outline the cross-section. While the islands on the left have small segments and provide a nice approximation, the islands on the right use the same number of segments that are larger, resulting in the segments overlapping and providing a poor approximation; (b) the same islands using smaller segments that better approximate the cross section.

Being code independent allows the analysis to deployed in multiple ways. Currently, it is deployed as part of the VisIt Visualization system, which has a rich set tools for parallel and remote processing. For single-domain data, parallelizing the integration and analysis on a per fieldline basis is trivial. As the number of domains grows, however, the integration of the fieldlines becomes more complex. To address this problem, we utilized a hybrid parallelization technique that dynamically adapts the strategies based on processor utilization. With this hybrid strategy a balance between redundant I/O and fieldline communication is sought that maximizes processor utilization. Complete details can be found in [6].

The analysis also can be run in situ utilizing the simulation code's parallelization and integration schemes and the finite elements directly. This latter usage cannot be overemphasized as the higher-order finite elements used within the fusion community are unique to the simulation code and must be replicated within VisIt and other postprocessing tools to ensure accurate analysis.

3. Identifying "Interesting" Particles Using Query-Based Tools

Another area of concern for scientists designing fusion reactors is a phenomena known as microturbulence, which plays a critical role in energy transport. Microturbulence causes heat to be transported from the hot plasma core to the outer walls of the reactor, thus degrading the confinement. One way to model microturbulence is by using particle-in-cell (PIC) codes that follow the trajectory of millions to trillions of individual particles to simulate the plasma [4]. Storing multivariate data for such a large quantity of particles can be challenging since scientists cannot predict in advance which particles will contribute the most to the transport of energy.

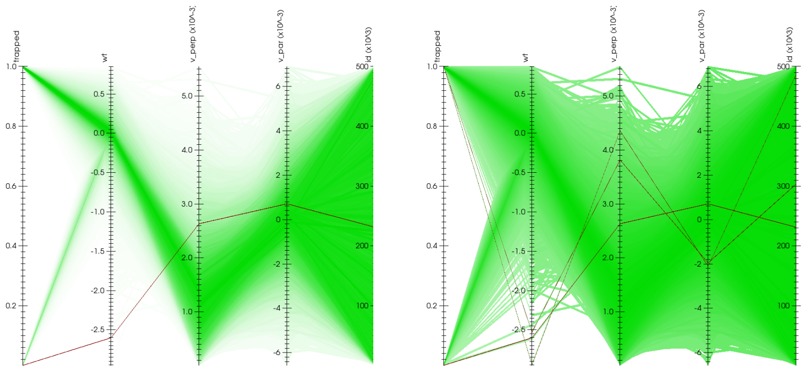
Therefore scientists are seeking ways to explore their data and identify particles of interest via their multivariate signature. To address this need, we have developed a tool that combines parallel coordinates [1] with fast index searches [11] to allow interactive query-based visualization and analysis. We briefly describe the system, which is still being developed.

The parallel coordinates method is a rich and power way to look at multidimensional data [1]. Each variable is represented by a line, X_i, in a 2D. A vertex on the line represents a single data value. In order to represent multivariate data point $(c_0, c_1, c_2, ..., c_n)$, a polygonal line with segments (c_0, c_1), (c_1, c_2), ... (c_{n-1}, c_n) is drawn. Figure 4a shows one such data point (the polygonal line drawn in brown) for a five-dimensional piece of data. This step is done for each data point in the set and is reflected in Figure

4 by the series of green lines, which for many data values becomes very dense.

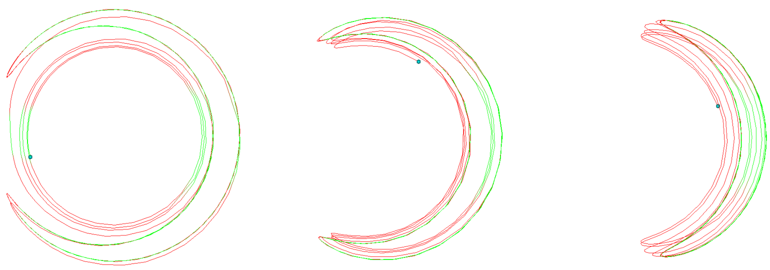
In order to address the density issue, the opacity of each line segment is adjusted based on the frequency that it would be drawn. In essence, a 2D histogram is constructed between each adjacent axes. The greater the frequency, the more opaque the line segment. For instance, in Figure 4a, there does not appear to be any line segments (i.e., data) for weights less than -0.5. Yet we were able to highlight at least one such data point. The reason is that the vast majority of the weights are near zero. As such, those values are more opaque, whereas the less frequent values are almost fully translucent. Figure 4b shows the same plot as in Figure 4a but with less contrast and shows the line segments for all the data values including the weights. More details on this visualization technique for data exploration can be found in [7].

One of the simplest ways to explore the data is by simply placing bounds on the axis and look for trends in the data. For instance, in Figure 4b using the last time step for a GTC simulation [4] involving 0.5 million particles, the bounds for the statistical weights were set to be smaller than -2.5 (the weight relates mainly to the particle's radial excursion from its original position). This query resulted in three particles being selected, drawn in brown in Figure 4b. Each of the particles has similar signatures with the exception of their passing/trapped state. Two of the three particles are in a magnetically trapped state, while the remaining one is in a passing state. The question for the scientist is whether this difference in state has any relevant meaning. In this case it does not seem to, ar



(*a*) (*b*)

Figure 4. (a) Parallel coordinate plot highlighting a single multivariate value (brown). In green are all of the particle multivarite values using the opacity mapping. (b) Parallel coordinate plot highlighting three multivariate values (brown) that have the largest radial excursions as bounded by the second axis. In green are all of the particle multivarite values with minimal opacity mapping.



(*a*) (*b*) (*c*)

Figure 5. Radial paths of three particles that exhibited the largest weights for the last time step in a GTC simulation. The paths are colored based on the particle's state, passing (green) or trapped (red).

However, if the radial (R-Z) paths for each of the three particles are drawn for the entire simulation there is a significant difference. For two of the particles, even though they have a large radial excursion, their paths follow a similar trajectory in that they have the characteristic "banana-shaped" orbit of magnetically trapped particles, as seen in Figures 5b and 5c. The particle in 5a, however, has been caught in the microturbulence eddies; and in addition to having a large radial excursion it was pushed through multiple trapped and passing orbits, as demonstrated by the path it followed during the simulation.

This is just one of many *what if* questions fusion scientists can explore as part of acquiring an understanding of their data. in order to ask such questions, the process must be interactive. Underlying the *what if* queries is the utilization of the FastBit software [11], which is a set of compressed bitmap indexes. These indexes provide efficient searching and retrieval operations and are used to accelerate data accesses and reduce the query response time. The drawback is that the indexes must be created after the data has been processed; and, though compressed, they can be very large. The advantage is that through such exploration scientists can begin to learn the multivariate signature of different particles and can use it to further their understanding of microturbulence. Furthermore, if a multivariate signature is known with some certainty, scientists can use it to selectively store particles, thus greatly reducing their I/O requirements; or, they can use the information to combine the particle data into density functions, thus reducing the overall I/O requirements.

4. Conclusions

We have presented two tools for data exploration. Both tools have been deployed as part of the VisIt software and are freely available to scientists. Both tools help provide scientists insight into their understanding of magnetic confinement fusion. While the tools have primarily been used for postprocessing, they are designed to aid scientists as they undertake larger simulations that will require in situ analysis as well reducing their I/O needs.

Both tools are works in progress. As scientists move toward computational steering, they will want to be able to track the growth of magnetic islands. Moreover, when microturbulence becomes too pronounced, scientists will also want to be able to track such regions. We envision both being incorporated into the simulation process, albeit in different forms.

Acknowledgments

We acknowledge the many scientists involved with the following DOE SciDACs: Visualization and Analysis in Support of Fusion Science SAP, Visualization and Analytics Center for Enabling Technology, Scientific Data Management Center for Enabling Technology, Center for Extended Magnetohydrodynamic Modeling, Center for Gyrokinetic Particle Simulations of Turbulent Transport in Burning Plasmas, and Center for Plasma Edge Simulation.

References

[1] Alfred Inselberg. *Parallel Coordinates*. Springer-Verlag, 2009.

[2] S.C. Jardin, J. Breslau, and N. Ferraro. A high-order implicit finite element method for integrating the two-fluid magnetohydrodynamic equations in two dimensions. *Journal of Computational Physics*, 226(2):2146 – 2174, 2007.

[3] S.E. Kruger, D.D. Schnack, and C. R. Sovinec. Dynamics of the major disruption of a $diii-d$ plasma. *Phys. Plasmas*, 12:56113, 2005.

[4] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and R. B. White. Turbulent transport reduction by zonal flows: Massively parallel simulations. *Science*, 281(5384):1835–1837, 1998.

[5] P. J. Morrison. Magnetic field lines, hamiltonian dynamics, and nontwist systems. *Phys. Plasmas*, 7:2279, 2000.

[6] D. Pugmire, H. Childs, C. Garth, S. Ahern, and G. Weber. Scalable computation of streamlines on very large datasets. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, November 2009.

[7] Oliver Rbel, Prabhat, Kesheng Wu, Hank Childs, Jeremy Meredith, Cameron G. R. Geddes, Estelle Cormier-Michel, Sean Ahern, Gunther H. weber, Peter Messmer, Hans Hagen, Bernd Hamann, and E. Wes Bethel. High performance multivariate visual data exploration for extemely large data. In *SuperComputing 2008 (SC08)*, 2008.

[8] Allen Sanderson, Guoning Chen, Xavier Tricoche, and Elaine Cohen. Understanding quasi-periodic fieldlines and their topology in toroidal magnetic fields. *Topological Methods in Data Analysis and Visualization II – Theory, Algorithms, and Applications*, 2011.

[9] Allen Sanderson, Guoning Chen, Xavier Tricoche, David Pugmire, Scott Kruger, and Joshua Breslau. Analysis of recurrent patterns in toroidal magnetic fields. *IEEE Transactions on Visualization and Computer Graphics*, 16:1431–1440, 2010.

[10] C.R. Sovinec, D.C. Barnes, T.A. Gianakon, A.H. Glasser, R.A. Nebel, S.E. Kruger, D.D. Schnack, S.J. Plimpton, A. Tarditi, M.S. Chu, and the NIMROD Team. Nonlinear magnetohydrodynamics with high-order finite elements. *Journal of Computational Physics*, 195:355, 2004.

[11] Kesheng Wu, Sean Ahern, E. Wes Bethel, Jacqueline Chen, Hank Childs, Estelle Cormier-Michel, Cameron Geddes, Junmin Gu, Hans Hagen, Bernd Hamann, Wendy Koegler, Jerome Lauret, Jeremy Meredith, Peter Messmer, Ekow

Otoo, Victor Perevoztchikov, Arthur Poskanzer, Prabhat, Oliver Rubel, Arie Shoshani, Alexander Sim, Kurt Stockinger, Gunther Weber, and Wei-Ming Zhang. Fastbit: Interactively searching massive data. In *SciDAC 2009*, 2009.