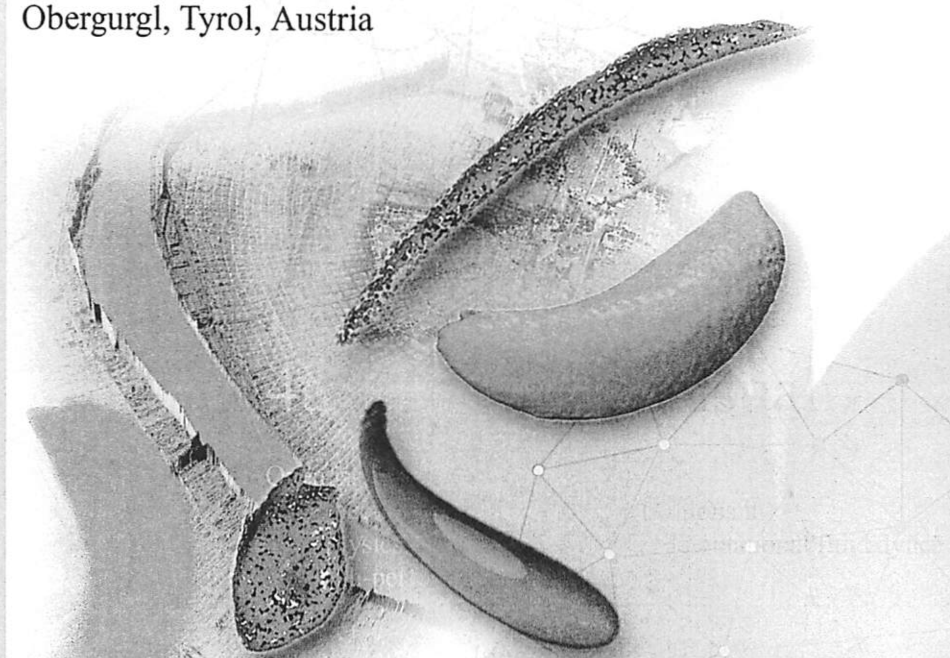
Werner Benger
René Heinzl
Wolfgang Kapferer
Wolfram Schoor
Mayank Tyagi
Shalini Venkataraman
Gunther H. Weber
(Eds.)

# Proceedings of the 4th High-End Visualization Workshop

Open issues in visualization
with special concentration on applications in
astrophysics, numerical relativity, computational fluid dynamics
and high-performance computing

June 18th- 22th, 2007
Obergurgl, Tyrol, Austria



http://www.vizworkshop.at

# A Visualization Toolkit for Lattice Quantum Chromodynamics

Massimo Di Pierro
School of Computer Science, Telecommunications and Information Systems
DePaul University, 243 S Wabash, Chicago, Illinois, USA
mdipierro@cs.depaul.edu

**Abstract**

Lattice QCD is an algorithmic formulation of QCD, the mathematical model that describes how quarks bind together to form composite particles such as proton and neutron. It has been successful in predicting properties of many newly discovered particles, including their mass and decay time. Unfortunately, lattice QCD is very computationally expensive and comprises of sophisticated algorithmic manipulations of large data-structures whose interpretation is purely statistical. In this paper we provide an overview of both lattice QCD and our work to develop a visualization toolkit to extract information from those data-structures. Our toolkit consists of a set of parallel algorithms for projecting the lattice QCD data structures into 3D scalar fields (for example the topological charge of the vacuum, the energy density, the wave function of the quarks, etc.) and uses VTK for the proper visualization.

## 1 Introduction

Quantum Chromodynamics [Marciano & Pagels, 1978] (QCD) is the mathematical model that best describes interactions among quarks, the basic constituents of most of ordinary matter. Lattice QCD is a formulation of QCD in terms of discretized space and time (lattice) that is suitable for numerical computation. Lattice QCD has been successful at explaining and predicting properties of composite particles such as the mass and lifetime of protons, neutrons, and many other particles produced by modern particle accelerators such as the Tevatron [FERMILAB-Pub-01/197, ] at Fermilab, LEP and LHC at CERN, and Slac at Stanford.

For a compact introduction to lattice QCD see [Pierro, 2006] and references therein.

Lattice QCD algorithms comprise of massive parallel Monte Carlo simulations. Modern state of the art computations are performed on commercial supercomputers (such as Blue Gene [Bhanot et al., 2005] and the Earth Simulator), clusters of workstations (there are nearly 1000 dedicated nodes at Fermilab [Holmgren, 2005] and Jefferson Laboratory), and dedicated machines (Ape [apeNEXT Collaboration, 2003] in Rome, QCDOC [Gottilieb, 2006] at Columbia University and Brookhaven National Laboratory).

In many Lattice QCD computations, the only published output consists of one number with two or three significative digits. At the same time, hundreds of terabytes of data are generated in the intermediate steps of the computation and are not usually looked at because their interpretation is purely statistical: they are random points in a Monte Carlo ensemble.

The goal of our project is twofold: identifying a set of projection operators that would map this data into 3D scalar fields of physical significance for the purpose of extracting information in a visual format; incorporating these operators into a visualization toolkit that will interface with existing Lattice QCD software libraries such as MILC [Gottlieb, 2002], FermiQCD [Pierro, 2002] and SciDAC [Brower, 2006].

In the next section we will give a brief description of Lattice QCD from a computational point of view and we will discuss examples of projection operators of physical interest. In the third section we will present the high level design of our toolkit. Finally we will draw conclusions, show some images produced by our system and discuss the current status of the project.

## 2 Theoretical Foundations

The ingredients of any Lattice QCD computation are the following:

- In nature, there are 6 *flavors* of *Quarks*. They are the basic constituents of protons and neutrons, as well as of other composite particles that can be produced in particle accelerators and are occasionally produced naturally by cosmic rays. Each quark flavor can best be described as a complex field $q_{x,\alpha,i}$. The index $x = (\vec{x}, t)$ labels a point on the hypercubic lattice that corresponds to a position in space $\vec{x}$ and time $t$. At point $x$ in space-time, a quark exhibits a number of local degrees of freedom that are parameterized by the indices $\alpha = 0, 1, 2, 3$ and $i = 0, 1, 2$. $\alpha$ is referred to as spin index and $i$ as color index. $|q_{x,\alpha,i}|^2$ is the probability of finding the quark at a given position $\vec{x}$ at time $t$, in a spin state $\alpha$ and color state $i$.

- Quarks interact with each other by exchanging *gauge bosons*, also known as *gluons*. Gauge bosons can be best described as a complex field $U_{x,\mu,i,j}$. Similarly to the case of quarks, the index $x$ labels a point in space-time, while the indices $\mu$, $i$ and $j$ parameterize the local degrees of freedom. If we define $P_{x,\mu,\nu} = U_{x,\mu}U_{x+\mu,\nu}U^\dagger_{x+\nu,\mu}U^\dagger_{x,\nu}$ and $U_{x,-\mu} = U^\dagger_{x-\mu,\mu}$ then[1]

$$F^a_{x,\mu,\nu} = \frac{1}{8} Re\ Tr[\lambda^a(P_{x,\mu,\nu} + P_{x,-\mu,\nu} + P_{x,-\mu,\nu} + P_{x,-\mu,-\nu} \quad (1)$$

$$- P^\dagger_{x,\mu,\nu} - P^\dagger_{x,-\mu,\nu} - P^\dagger_{x,-\mu,\nu} - P^\dagger_{x,-\mu,-\nu})] \quad (2)$$

  is the chromo-electro-magetic tensor associated to gluons of type $a$. For each gluon type, $E_{x,i} = F_{x,0,i}$ is the chromo-electric field and $B_{x,k} = \sum_{i,j} \epsilon_{i,j,k} F_{x,i,j}$ is the chromo-magnetic field.

- QCD is a *Quantum Field Theory*. This means that there is no deterministic time-evolution for the above fields. In fact, the only meaningful physical observables in any Quantum Field Theory are the *correlations* between the degrees of freedom. Lattice QCD provides prescription rules on how to measure physical observables (for example the mass of a proton) using correlations among field variables. These correlations are computed numerically by averaging the corresponding operator over multiple *field configurations*, also known as *paths* or *histories*. Each configuration represents a possible evolution in time of a small portion of space of about $(10^{-14}$ meters$)^3$ for about $10^{-22}$ seconds.

- Field configurations are generated via a Markov Chain Monte Carlo (MCMC) algorithm using a transition probability that encodes the physical laws of QCD. The "Quantum" aspect of the theory is represented by the stochastic field fluctuations present in the gauge configurations and averaged over.

[1] here $\lambda^a$ is any of the 8 generators of the $SU(3)$ group

- For practical purposes any Lattice QCD computation is divided into three main steps: 1) gauge field configurations are generated using the MCMC; 2) for each gauge configuration one places the quarks in a certain state and let them evolve according to the Dirac equation in the background gauge field (the solution of the Dirac equation on a given gauge configuration is called a *quark propagator*); 3) the indices of a gauge configuration and its corresponding quark propagators are contracted together to compute the operator corresponding to a given physical observable, which is then averaged over all the gauge configurations.

- A typical gauge configuration $U$ has a size of 96 points in time and $64^3$ points in space, corresponding to 7 gigabytes of data. A typical quark propagator $q$ for a single source on the above gauge configuration has a size of 2.5 gigabytes. Most Lattice QCD computations involve about 1000 gauge configurations and a minimum of 12 quark sources each, thus generating $10 \div 100$ terabytes of data. Usually gauge configurations and quark propagators are stored and are reused for multiple observables in a semi-industrial fashion. Typical computations require 1-10 million hours of computing time for a Pentium 4 @ 3GHz equivalent CPU.

- One complication consists in the fact that some of the degrees of freedom in the gauge field $U$ are redundant but cannot be eliminated in the computation. In fact $U_{x,\mu,i,j}$ must be unitary matrices in the indices $i,j$ and only operators invariant under the simultaneous transformations $U_{x,\mu,i,j} \rightarrow \sum_{m,n} \Lambda_{x,i,m} \Lambda^*_{x+\mu,j,n} U_{x,\mu,m,n}$ and $q_{x,\alpha,i} \rightarrow \sum_m \Lambda_{x,i,m} q_{x,\alpha,m}$ have a physical significance[2]. The above symmetry is called *gauge invariance* and it puts a major constraint on what can be visualized. The gauge invariance symmetry represents the principle of local indistiguishibility among quarks of different colors (they are 3 but one cannot tell them apart). This symmetry, motivated by experiments, poses a strong contraint on the model and it almost completely determines the inteaction term between quarks and gluons in the Dirac equation for QCD.

- The only input of a Lattice QCD computation are parameters that

[2]Here $\Lambda_{x,i,j}$ is an arbitrary field of unitary matrices in the indices $i,j$

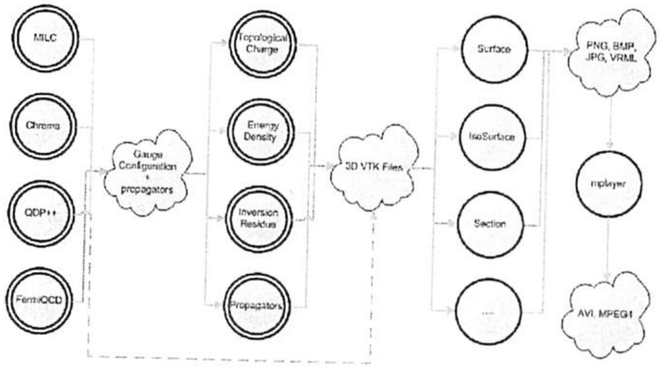



Figure 1: The figure shows the data flow in our visualization toolkit. The double circles represent parallel components including Lattice QCD algorithms and projection plug-ins. The single circle represent Mayavi visualization plug-ins.

loosely correspond to the masses of the quarks and to the lattice discretization scale. Physical observables have a dependence on the discretization scale that is known as *running*. For a sufficiently small discretization scale, the effect of the running can be corrected for and the results of the computation, in physical units, can be made independent on it. Physical observables also depend on the quark masses and, in fact, these are currently determined by a fine tuning the input parameters in Lattice QCD computations.

There are multiple reasons why extracting visual information from Lattice QCD data is important and here we list some of them:

- The gauge configurations have a non-trivial topological structure that can be isolated by removing short term fluctuations, a process called *cooling*. It is known that the total topological charge changes very slowly under MCMC steps, but it is not know how the local charge evolves.
- Some observables are themselves extended objects (for example the wave function of a quark inside a hadron or the energy density in pres-

ence of a quark-antiquark couple) but so far only 1D or 2D sections are usually visualized.

- The algorithms used for the MCMC and to invert the Dirac operator are both iterative. Therefore it is interesting to visualize how information propagates, step by step, across the lattice in order to understand the local convergence of these algorithms.

- Because of the size of the data structures involved, Lattice QCD algorithms are implemented as tightly-coupled parallel programs written in C/C++. Visualization can help monitor the communication patterns and identify bugs and network problems.

- Lattice QCD requires knowledge of multiple disciplines and therefore it has a very steep learning curve. The visualization of actual computations can serve a didactic purpose thus fostering a better intuitive understanding of QCD and pushing scientific progress.

## 3 Implementation

The main architecture of our system, fig. 1, is comprised of two main parts: a set of projection operators implemented in C/C++ as parallel MPI programs and a graphical interface based on the Enthough Workbench [wor, 2007] and Mayavi 2.0 [Ramachandran, 2001], which implements a Python interface to VTK [Schroeder et al., 2000]. We will refer to the former as a projection plug-in as opposed to the visualization plug-ins provided by Mayavi.

An example of a projection plug-in is a parallel algorithm that reads gauge configurations, cools it, computes the topological charge in 4 dimensions, takes a time slice and saves it as 3D scalar field in the VTK file format.

An example of a visualization plug-in is an algorithm that reads the above VTK file and generates iso-surfaces.

The Workbench, fig. 2, provides a GUI for the entire system and allows users to interactively manipulate the VTK files: display, rotate, edit, save them in some other standard graphical formats, including JPG, PNG, and VRML.

Mayavi is a Python interface to VTK and allows scripting of the above operations. A typical script would loop over a large set of fields, process each

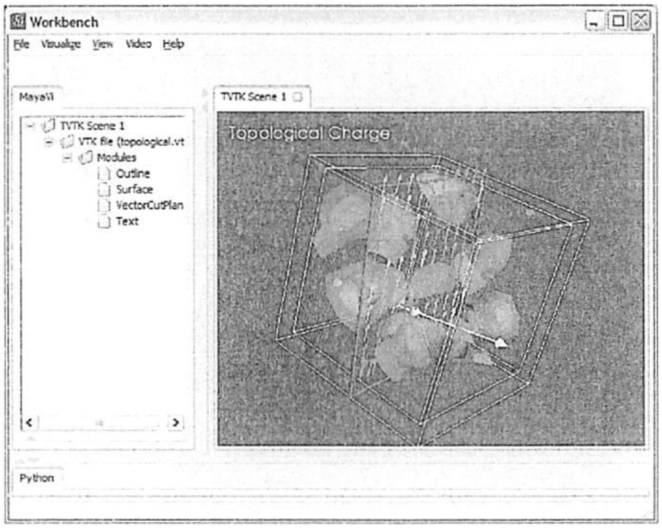



Figure 2: A screenshot of the Enthought Workbench showing the topological charge for a time-slice of a 4D gauge configuration.

of them using the same plug-ins and produce the individual frames as an animation.

Independently on the set of plug-ins used we identified three general recurrent patterns:

- Given one configuration, project and visualize the different time-slices.
- Given one set of configurations and one time-slice $t$, project and visualize the same time slice for each configuration.
- Given a set of configurations, for each time-slice, project the time-slice on each configuration, average over all configurations and visualize the average time-slice.

In order to remove visual unpleasant effects of high-frequency quantum fluctuations, when necessary, we adopted the following gauge-invariant smooth-

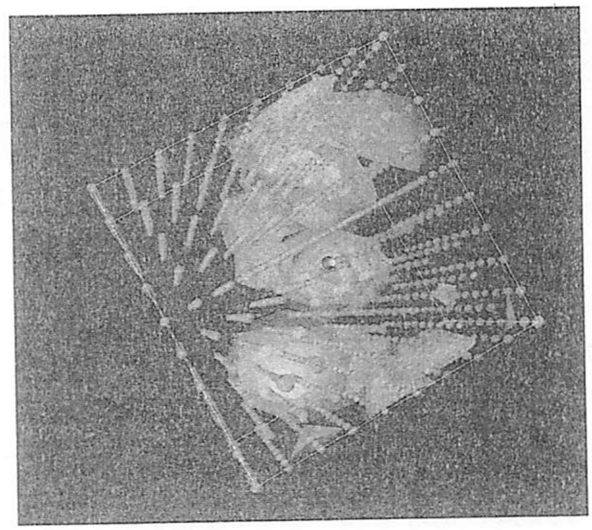


Figure 3: The figure shows iso-surfaces for the topological charge for a time-slice of small test gauge configuration. The spheres show the lattice geometry.

ing algorithm for the gauge configuration:

$$U_{x,\mu,i,j} = \mathcal{P}[\xi U_{x,\mu,i,j} + (1-\xi)\sum_{n,m} U_{x,\nu,i,m} U_{x+\nu,\mu,m,n} U^*_{x+\mu,\nu,j,n} \tag{3}$$

$$+ U^*_{x-\nu,\nu,m,i} U_{x-\nu,\mu,m,n} U_{x+\mu-\nu,\nu,n,j}] \tag{4}$$

Here $x \pm \mu$ are the coordinates of a lattice point shifted from $x$ in direction $\pm\mu$, $\xi$ is an arbitrary smoothing parameter, and $\mathcal{P}$ is a projection operator into the $SU(3)$ group.

## 4 Examples and Conclusions

This project started about six months ago. We have successfully produced a set of prototype projection plug-ins that compute topological charge, wave functions, energy density, and density of heatbath hits. Examples of images are shown in figs. 3 and 4.

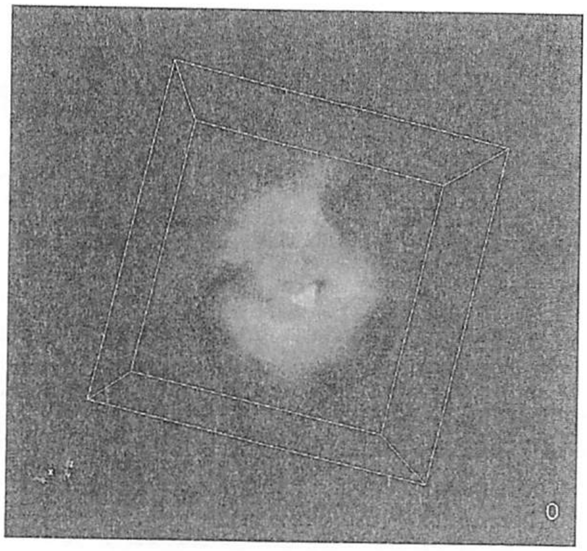


Figure 4: Wave function of a heavy-light meson computed on a single gauge configuration.

The field of visualization of Lattice QCD data is still in its infancy, but has a large potential impact for the physics community. It will help us better understand local properties of the algorithms and the spatial characteristics of extended physical objects such as mesons, hadrons and glue-balls. We also strongly believe that visualization is important to better explain what Lattice QCD is and to attract more students to this exciting area of research.

# 5 Acknowledgments

This project is supported by the Department of Energy under the Scientific Discovery Through Advanced Computing grant DE-FC02-06ER41441. We wish to thank the Fermilab theory group for continuing to share their data and their computing resources with the author. We also thank Dr. P. Ramachandran for developing Mayavi and his email support to this project. We finally thank the DePaul students involved in the project: Joel McGraw,

Tony Garcia, and Martin Hulth.

## References


[wor, 2007] (2007). Workbench. Available from: `http://www.enthought.com`.

[apeNEXT Collaboration, 2003] apeNEXT Collaboration (2003). The apenext project. *Nuclear Physics B - Proc. Suppl.*, 119, 1038–1040.

[Bhanot et al., 2005] Bhanot, G. et al. (2005). Qcd on the bluegene/l supercomputer. *Nuclear Physics B - Proc. Suppl.*, 140, 823–825.

[Brower, 2006] Brower, R. (2006). National software infrastructure for lattice quantum chromodynamics. *Journal of Physics - Conf. Series*, (pp. 142).

[FERMILAB-Pub-01/197, ] FERMILAB-Pub-01/197. *B Physics at the Tevatron: Run II and Beyond.* Technical report.

[Gottilieb, 2006] Gottilieb, S. (2006). Guest editor's introduction: Special purpose computers. *IEEE: Computing in Science and Engineering*, 8, 15.

[Gottlieb, 2002] Gottlieb, S. (2002). *Nuclear Physics B - Proc. Suppl.*, 106-107, 1031–1033.

[Holmgren, 2005] Holmgren, D. J. (2005). Pc clusters for lattice qcd. *Nuclear Physics B - Proc. Suppl.*, 140, 183–189.

[Marciano & Pagels, 1978] Marciano, W. & Pagels, H. (1978). Quantum chromodynamics. *Physics Reports*, 36, 135.

[Pierro, 2002] Pierro, M. D. (2002). Fermiqcd: A toolkit for parallel lattice qcd applications. *Nuclear Physics B - Proc. Suppl.*, 106-107, 1034–1036.

[Pierro, 2006] Pierro, M. D. (2006). An algotihmic approach to quantm field theory. *International Journal of Modern Physics A*, 21(3).

[Ramachandran, 2001] Ramachandran, P. (2001). Technical report.

[Schroeder et al., 2000] Schroeder, W. J., Avila, L. S., & Hoffman, W. (2000). Visualizing with vtk: A tutorial. *IEEE Computer Graphics and Applications*, 20(5), 20–27.