ARONE: ADAPTIVE REAL-TIME OPTIMIZATION FOR NEUTRON SCATTERING EXPERIMENTS

Janey Duong
University of California, Berkeley
Research Alliance in Math and Science
Computational Sciences and Engineering Division, Oak Ridge National Laboratory
Mentor: Yu (Cathy) Jiao

http://sites.google.com/a/g.ornl.gov/janey-duong/

Introduction

Many of the developments and improvements of products and technologies in our everyday life involve neutron scattering researches such as the extending of laptop computer batteries' life, the effective and safe materials that are used to make bridges, engine parts, and airplanes, and the production of high quality plastic products. Though playing a significant role and costing high expense in the research community, when it comes to study in situ dynamic or kinematic behaviors of materials, researchers do not have a tool to predict where and when the interesting phenomenon will occur yet. Most of the time, neutron scattering experiments are heavily dependent on scientists' initial guesses, and these guesses remain unchanged during the experiments. To help solving this problem, the Computational Sciences and Engineering Division proposes to develop a software tool that is capable of producing experimental plans adaptive to the measurements in real time. The research group aims to make this new tool work for at least three different types of non-linear functions of arbitrary shapes. The ultimate goals are to design and to implement an intuitive graphical interface that will allow users to use the algorithms. The result from this project will be essential for future researches involving neutron scattering science.

Background information

❖ Why is this tool needed?

Incorrect initial guesses lead to wasted:

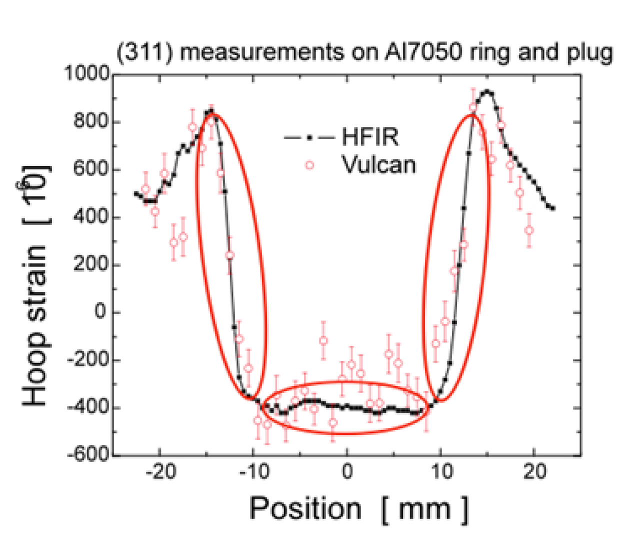
- ○ Time
- ○ Money

❖ Programming languages and software used:

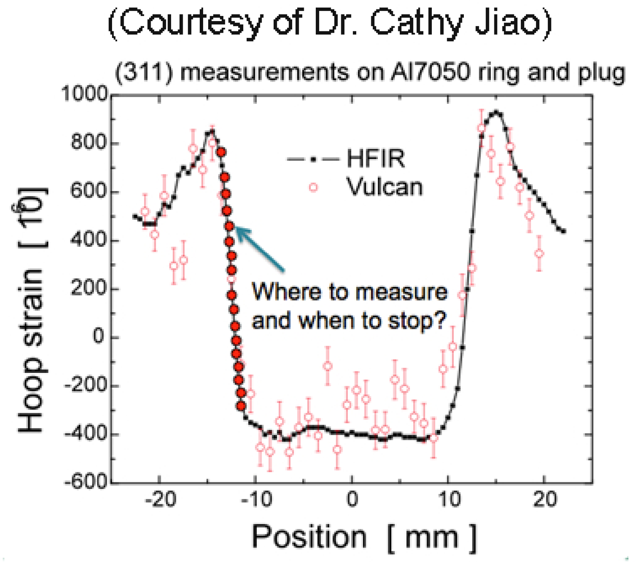
- ▪ Python: NumPy, SciPy, Matplotlib, Wxpython
- ▪ FFmpeg: to record, convert and stream audio and video
- ▪ Statistical language R

❖ Hypothesis testing:

- ▪ Kolmogorov-Smirnov Test
- ▪ Wilcoxon Test
  - ○ Random sampling
  - ○ Independent samples and measurements
  - ○ Non-parametric
  - ○ Two independent samples' comparison



(Courtesy of Dr. Cathy Jiao)



(Courtesy of Dr. Cathy Jiao)

Methods

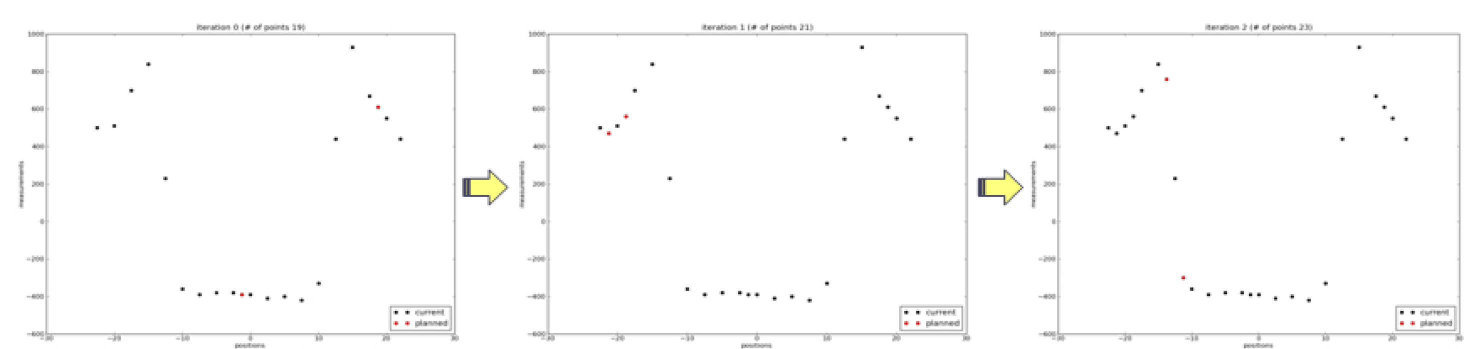
❖ Test case 1: demonstrate initial tool on an actual data

- ▪ Clean data: actual experimental data
- ▪ Noisy data: random Gaussian noise added to each point in the clean data

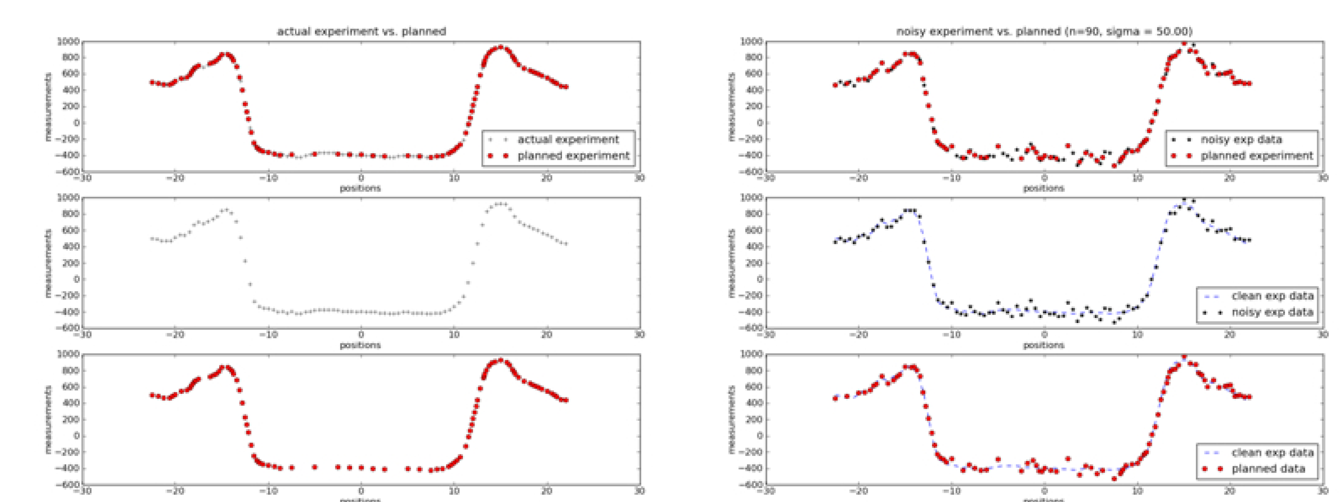
Planning scheme:

- ○ Choose a set of initial points based on range of the positions
- ○ Obtain the measurements for the chosen initial points
- ○ Compare the absolute changes in measurements between two neighboring points
- ○ For each iteration, add two midpoints in two intervals with the largest absolute changes until total sampling points met.

First three iterations of clean data solution



Final comparison plots



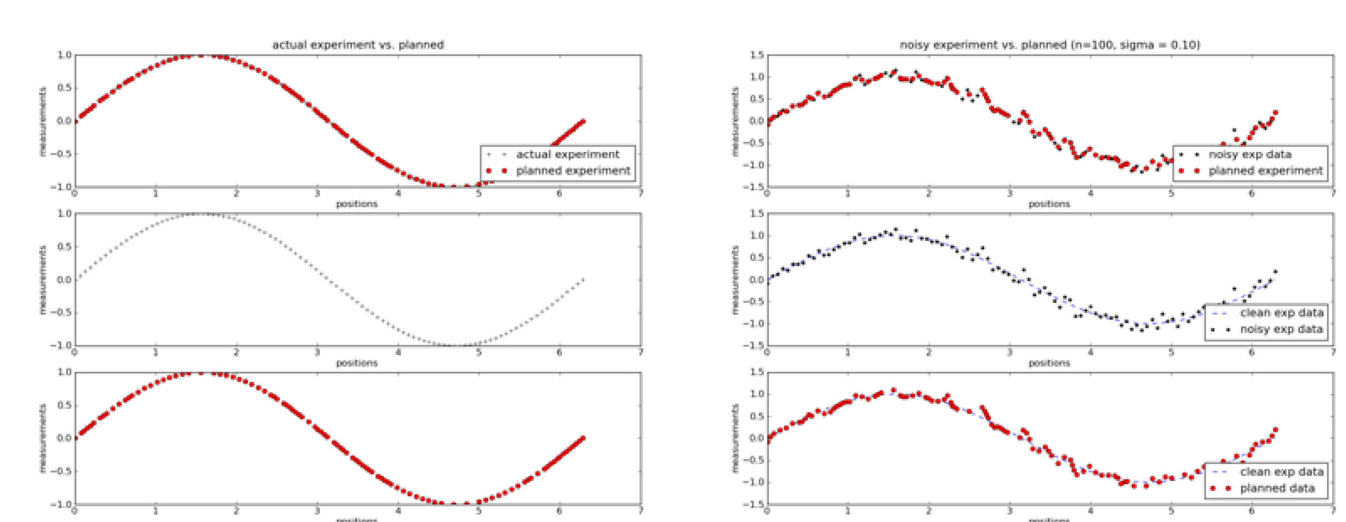
Clean data solution / Noisy data solution

⇒ Clean data solution: planned and experimental data graphs match each other.

⇒ Noisy data solution: assuming clean data to be ground truth, comparison plots show that noisy and planned data match each other.

❖ Test case 2: test/modify tool on sine function curve

Final comparison plots



Clean data sine / Planned data sine

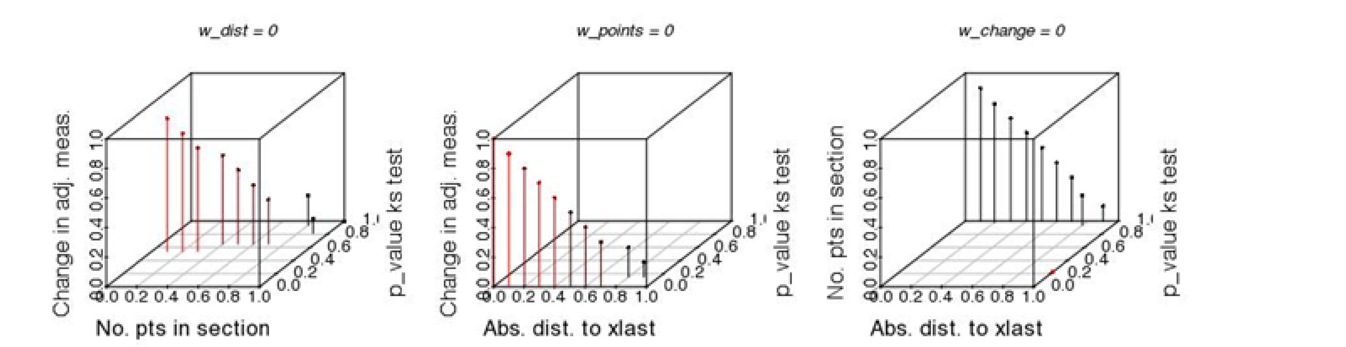
⇒ Clean data sine: assuming an arbitrary experiment's measurements fall exactly on the sine curve, initial tool returns a planned data that closely follows the experimental one.

⇒ Noisy data sine: assuming clean data sine to be the ground truth, the comparison plots show that initial tool returns a planned data that follows the noisy one quite well.

❖ Hypothesis testing:

- ▪ Null hypothesis: planned and experimental measurements drawn from same distribution
- ▪ Alternative hypothesis: planned and experimental measurements drawn from different distributions
- ▪ Tests: Kolmogorov-Smirnov and Wilcoxon tests
  - ○ If p-value > 0.05 → Accept null hypothesis
  - ○ If p-value < 0.05 → Reject null hypothesis

❖ Sensitivity analysis: Clean data solution



⇒ The weight of number of points in the section is significant. When this weight is set to 0, the Kolmogorov-Smirnov test shows that 8 out of 10 cases, initial tool fails to predict the measurement values (p-values < 0.05).

Test results

❖ Test case 1:

Kolmogorov-Smirnov test:	Wilcoxon test:
Clean data solution: p-value = 0.1087	Clean data solution: p-value = 0.3812
Noisy data solution: p-value = 0.9621	Noisy data solution: p-value = 0.7806

Accept null hypothesis

❖ Test case 2:

Kolmogorov-Smirnov test:	Wilcoxon test:
Clean data solution: p-value = 0.6557	Clean data solution: p-value = 0.9457
Noisy data solution: p-value = 0.7031	Noisy data solution: p-value = 0.832

Accept null hypothesis

Conclusion and future work

❖ Planned and experimental data derived from same distribution

- ▪ Initial tool works on an actual experimental data
- ▪ Initial tool works on sine function

❖ Movies were created from series of static images

❖ A graphical user interface was designed and implemented

❖ Future works:

- ▪ Test/modify tool on 2 additional non-linear functions
- ▪ Extend algorithm to test two dimensional data
- ▪ Make software tool available to users

References

Hatzivassiloglou, Vasileios. "The Kolmogorov-Smirnov Test." Lecture. *~vh/Courses. Web. 13 July 2010. <http://www.hlt.utdallas.edu/~vh/Courses/>*

"Mann-Whitney U Tests." Statistics Solutions. July 13, 2010 *<http://www.statisticssolutions.com/methods-chapter/statistical-tests/mann-whitney-u-test/>*

"ORNL Neutron Sciences." Oak Ridge National Laboratory. July 12, 2010 *<http://neutrons.ornl.gov/>*

The Research Alliance in Math and Science program is sponsored by the Office of Advanced Scientific Computing Research, U.S. Department of Energy. The work was performed at the Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC under Contract No. De-AC05-00OR22725. This work has been authored by a contractor of the U.S. Government, accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

I would like to thank my mentor, Dr. Cathy Jiao for guiding me through the project and helping me preparing for this presentation. Also, I would like to thank the Department of Energy, Oak Ridge National Laboratory, and the Research Alliance in Math and Science program for offering this great opportunity to work on a wonderful research project this summer.