Matthew D. Rocklin

CONTACT INFORMATION

E-mail: mrocklin@gmail.com *webpage:* http://matthewrocklin.com

RESEARCH INTERESTS

I build and maintain parallel computing infrastructure for Python's open source ecosystem. This is part of a broader effort to increase accessibility of numerical methods to science and policy practitioners.

EDUCATION

University of Chicago, Chicago, IL

Ph.D, Computer Science **2013**
M.S. Computer Science **2011**

University of California, Berkeley, Berkeley, CA

B.A., Physics, Mathematics, and Astronomy **May 2007**

PROFESSIONAL EXPERIENCE

Anaconda Inc
Computational Scientist **2014 - Present**
Parallel and distributed computation for the open source Python ecosystem

Sandia National Laboratory - Livermore, CA
Postdoctoral Researcher **2013 - 2014**
Computation and analysis of large time evolving networks

Sandia National Laboratory - Livermore, CA
Summer researcher **Summer 2010**
Complex networks, particularly with clustering and multiple metrics

Argonne National Laboratory - Chicago, IL
Givens Fellow **Summer 2009**
Uncertainty quantification and sensitivity analysis of numerical weather prediction

UC Berkeley Physics Department - Berkeley, CA
Staff Research Assistant **2007 - 2008**
Algorithms and software to probabilistically track intracellular movement of vesicles moving within the bodies of plant cells. Developed biophysics educational tools

Berkeley Engineering and Research/4D Imaging - Berkeley, CA
Developer **2003 - 2005**
3d scanner based on structured light techniques. Began startup engineering company. Initial sole developer of a project which eventually grew to become an independent and profitable company

SOFTWARE

I contribute to and maintain several libraries within Python's numeric computing ecosystem, particularly around parallel and distributed computing. I am primarily known for my current work on Dask.

- Dask – *Parallel computing with dynamic task scheduling*: Dask combines a high-speed computational task scheduler with algorithms for parallel arrays, dataframes, machine learning, etc..

However in the maintenance of the Dask project I also end up acting as steward over a large part of the numerics ecosystem, dealing with libraries for communication, data access, and so on.

Historically I have been involved in the following open source projects:

- SymPy – *Symbolic mathematics in Python*:
- Theano – *Mathematical array compiler for deep learning*:

A more complete list of software is available at github.com/mrocklin.

Publications

Today I mostly publish on technical topics on my blog at matthewrocklin.com/blog. Previoulsy I engaged in traditional academic publishing. This page contains references to those works.

Theses

- M. Rocklin, Modular Generation of Scientific Software, 2013, a PhD dissertation.
- M. Rocklin, Uncertainty Quantification and Sensitivity Analysis in Dynamical Systems, 2011, a masters thesis

Papers

- Al-Rfou, Rami, et al *Theano: A Python framework for fast computation of mathematical expressions*, arXiv preprint arXiv:1605.02688 (2016).
- A. Meurer et al *SymPy: symbolic computing in Python*, PeerJ Computer Science 3 (2017): e103.
- M. Rocklin, A. Pinar *On Clustering on Graphs with Multiple Edge Types*, Internet Mathematics, 2012
- E. Constantinescu,V. Zavala, M. Rocklin, S. Lee, and M. Anitescu, *A Computational Framework for Uncertainty Quantification and Stochastic Optimization in Unit Commitment with Wind Power Generation.* IEEE Transactions on Power Systems, 2010.

Conference Proceedings

- M. Rocklin, *Dask: Parallel computation with blocked algorithms and task scheduling*, Proceedings of the 14th Python in Science Conference. 2015.
- M. Rocklin, *Uncertainty Modeling with SymPy Stats* SciPy 2012
- M. Rocklin, A. Pinar, *Computing an Aggregate Edge-Weight Function for Clustering Graphs with Multiple Edge Types.* Algorithms and Models for the Web-Graph, 2010
- M. Rocklin, A. Pinar, *Latent Clustering on Graphs with Multiple Edge Types* Algorithms and Models for the Web-Graph, 2011

Other

- M. Rocklin, A Pinar, *Spectral Generation and Latent Community Structure of Multiweighted Networks*, 2010
- M. Rocklin, E. Constantinescu , *Adjoint Sensitivity Analysis for Wind Power Generation*, 2009
- US Patent 7620209: *Method and apparatus for dynamic space-time imaging system*