Data Analytics Activities

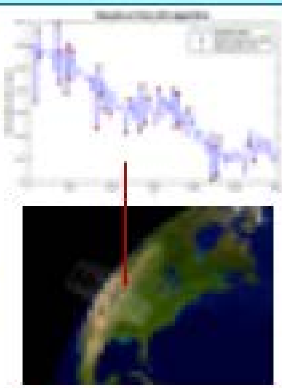
[1]Oak Ridge National Laboratory
[2]Dartmouth College
[3]Florida State University
[4]University of Tennessee Knoxville

FASTMath Team Members: R. Archibald[1], F. Bao[3], A. Gelb[2], H. Tran[1], & C. Webster[1&4]

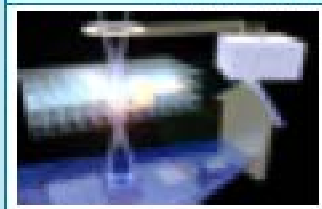
Sparse functional representation of data, to enable faster IO and analysis of big datasets

In-situ analysis of streaming climate data

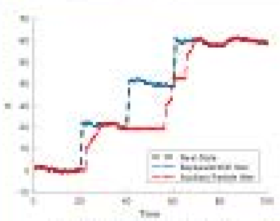
- Developing effective online data monitoring and archiving strategy over temporal and spatial domains while respecting practical storage and memory capacity constraints
- Streaming methods improves the quality and quantity of information available to scientists for analysis

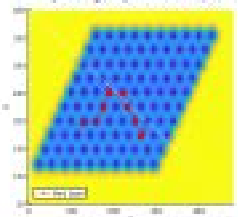


Atomic Forge Analysis



- Need fast accurate forward models for atomic states of tracked atoms.
- Need computational control of beam position and intensity.
- Will enable 3D atomic fabrication: quantum computing, spintronics, etc.





Left: Tracking of multiple well potential in 1D.
Right: Tracking of multiple well lattice potential in 2D

Sparse data representations

- We reconstruct data $c \in \mathbb{R}^{N\times l}$ from measurements $u \in \mathbb{R}^{m\times l}$ and $A \in \mathbb{R}^{m\times N}$:

$$u \approx Ac$$

  - Limited number of measurements: $m \ll N$.
  - The data are sparse.
  - $l = 1$: reconstructing a single dataset.
  - $l > 1$: simultaneously reconstructing multiple datasets.

- Recovery via regularizations enforcing sparsity:

$$c = \text{argmin}\ R(z) \ \text{subject to}\ u \approx Az$$

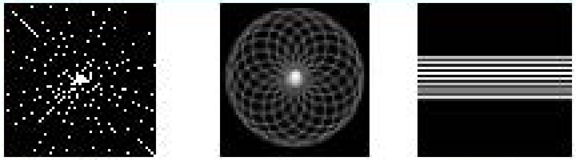
Standard CS: $R(z) = \|z\|_1$.
Structures of the sparsity can be exploited:

  - Downward closed and tree structures: $$R(z) = \|z\|_{\omega,1}.$$
  - Joint sparsity: $$R(z) = \|z\|_{2,1}.$$

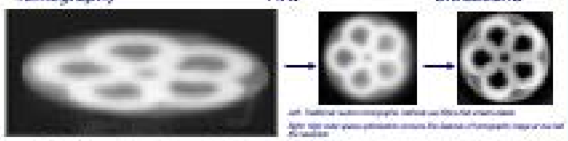
A comparison of adaptive weighted l_1 minimization with different choices of weights

$$R(z) = \|z\|_{\omega,1} \ \text{with}\ \omega_j = \max|A_{i,j}|$$

Sampling and fast algorithms



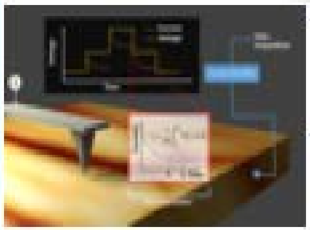
Tomography *MRI* *Ultrasound*



Thin plate projections of additive manufactured compound

Knowledge Extraction from Images

- Atomic resolution imaging data is readily abundant given advances in instrumentation
- No framework to extract knowledge from data
- We utilize concept of statistical distance to derive materials properties from atomic resolved data, providing a generative model exists



- Acquisition of current-voltage curves for electronic materials characterization at the nanoscale dates back decades
- We increase speed of acquisition by 1000x through full information capture + Bayesian inference methods

ion: http://www.fastmath-scidac.org or contact Rick Archibald, Archib