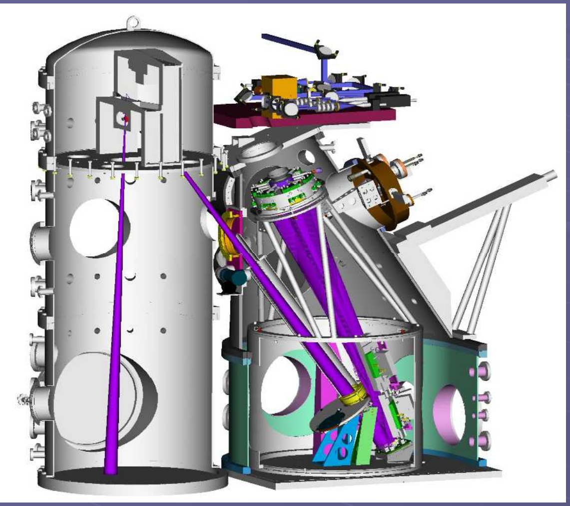
SAND2007-7820C

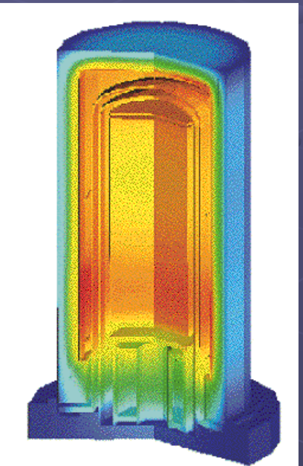
Developing a Data Extraction Toolkit for Optimization, Uncertainty Quantification, and Validation

Pamela J. Williams
Brian A. Adams
Genetha A. Gray
Sandia National Laboratories

Challenging optimization problems arise in a wide variety of applications

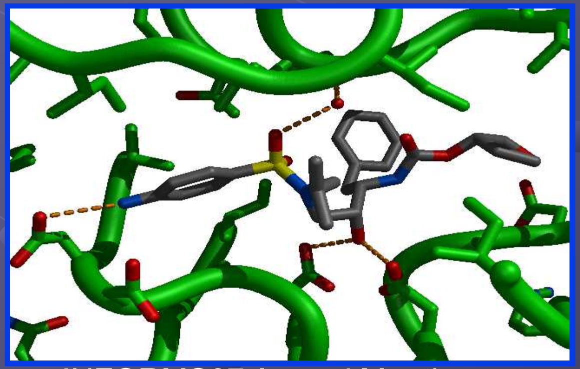


Parameter identification of an EUVL lamp model



Temperature fields in a vertical, stacked-wafer, low-pressure, chemical-vapor-deposition furnace

Optimal design and control of physical systems



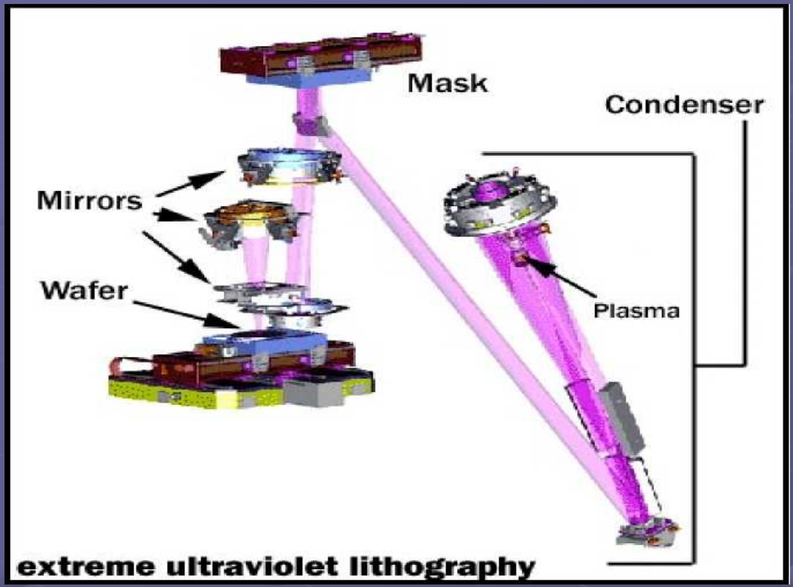
Determining minimum energy configurations in computational chemistry

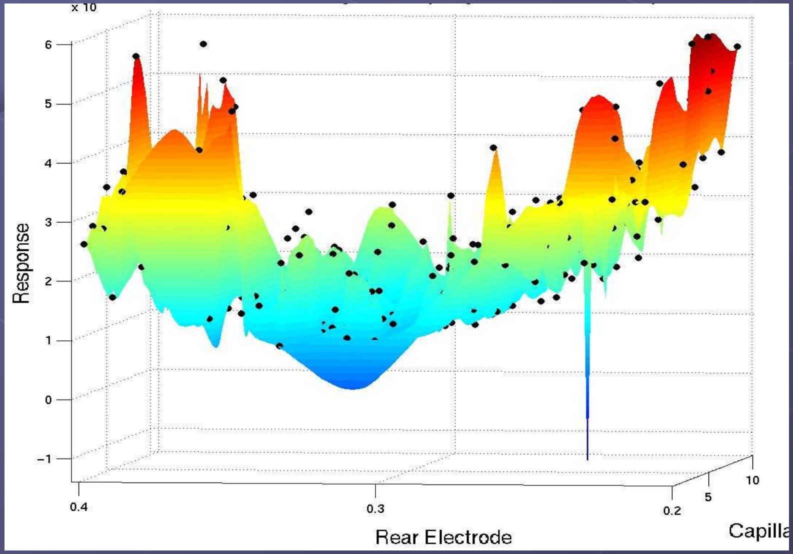
- Computer models characterized by
  - complex geometry
  - multiple materials
  - coupled sub-systems
  - complex physics
- Optimization problem based on simulation

Problem formulation

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s. t.} \quad & h(x) = 0, \\ & g(x) \geq 0 \end{aligned}$$

Optimization of an EUVL lamp model

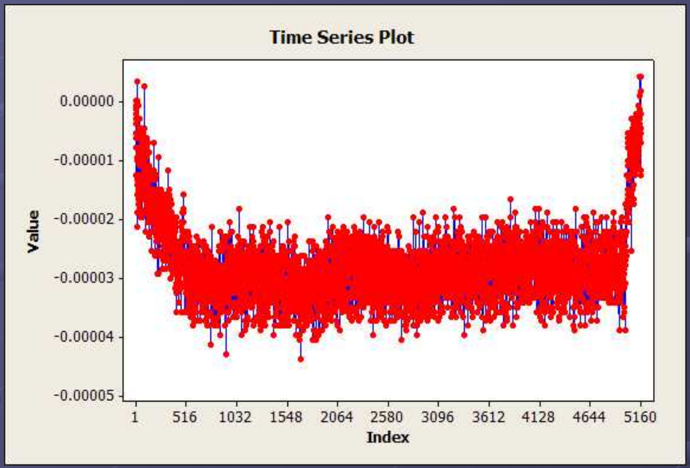


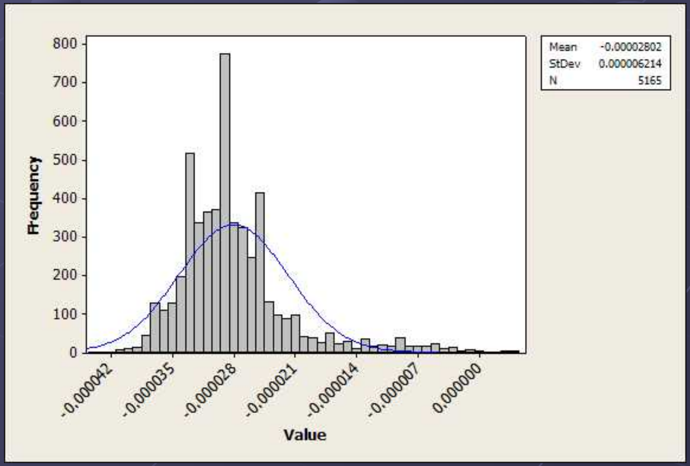


- *Problem:* identify thermal analysis model parameters which best match experimental data for an EUVL lamp
- *Formulation:* objective fn consisted of the maximum temperature difference over time histories for 6 probes.

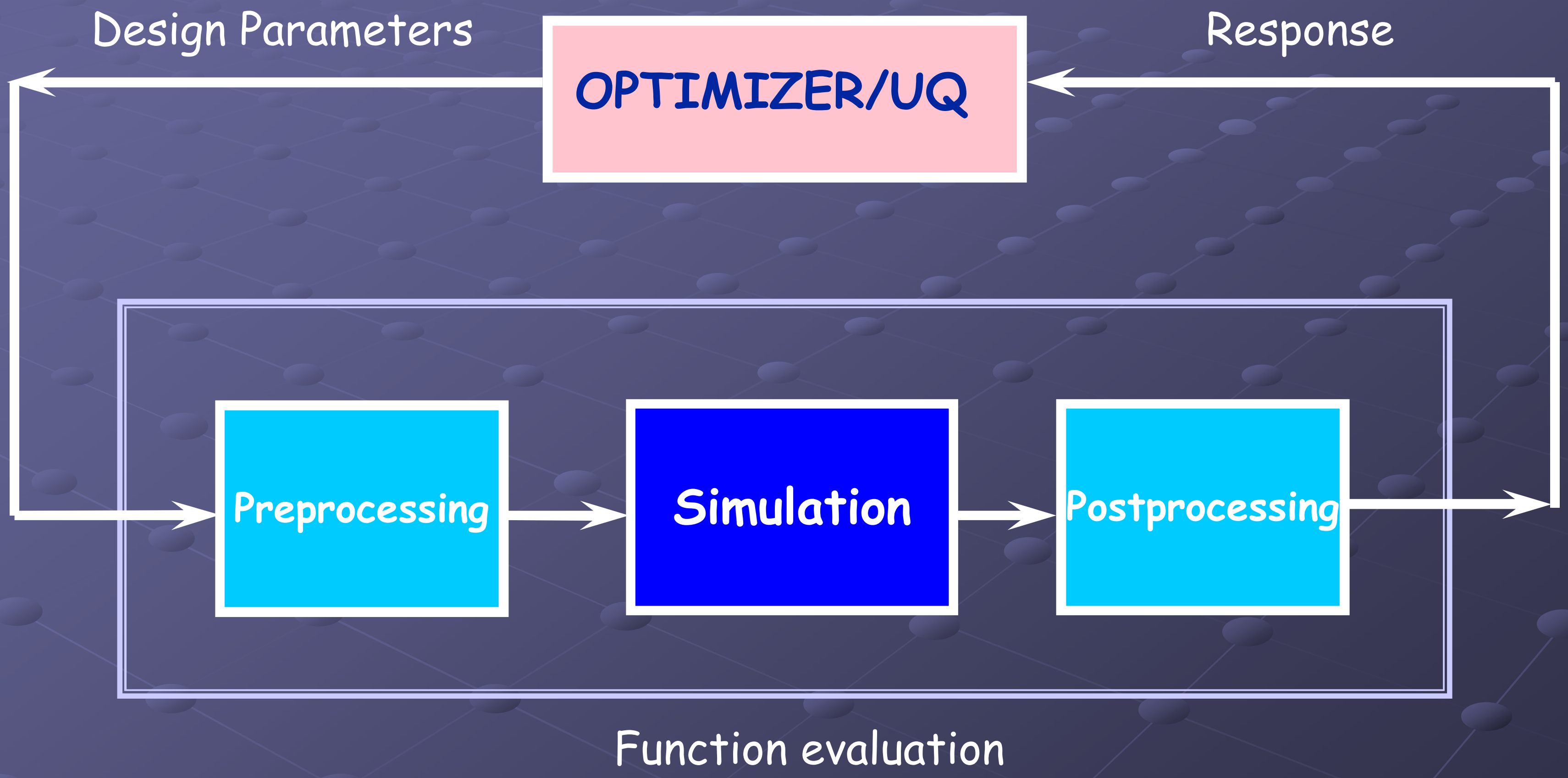
What is uncertainty quantification?

- Which parameters are "most important"?
- How sensitive is the simulation output to the model parameters?
- What is the mean and variance of simulation output?



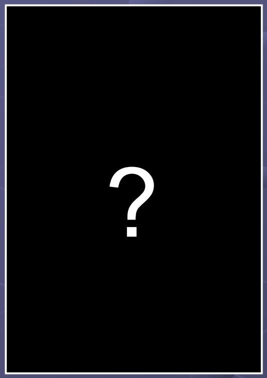


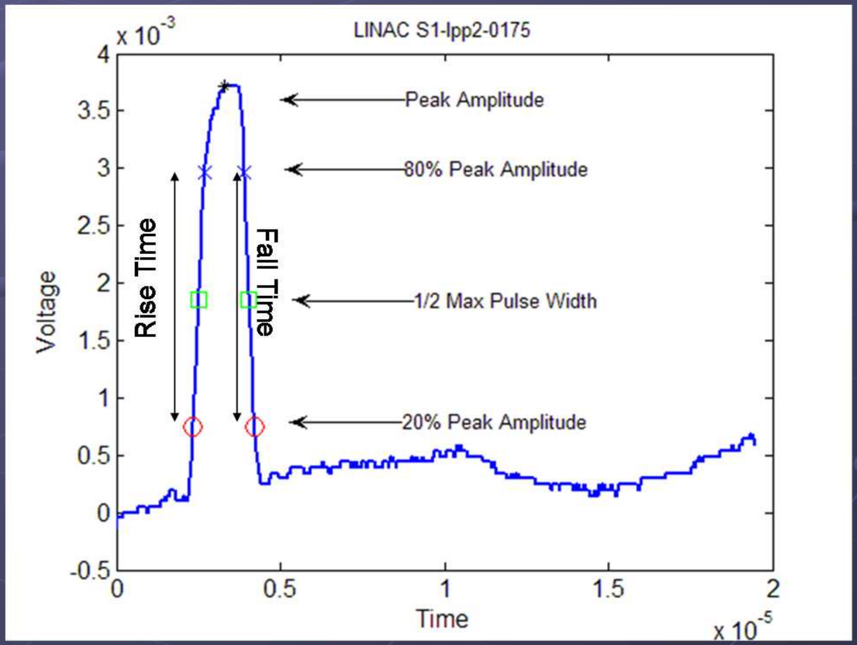
Optimization and UQ methods both need data



Given a time series dataset, compute signal characteristics

0.00E+00	2.52E-07
1.00E-09	-1.35E-06
2.00E-09	-6.15E-06
3.00E-09	-5.48E-07
4.00E-09	-2.15E-06
5.00E-09	-3.75E-06
6.00E-09	-5.35E-06
7.00E-09	-2.95E-06
8.00E-09	-7.75E-06
9.00E-09	-5.48E-07
1.00E-08	-1.25E-05
1.10E-08	-5.48E-07
1.20E-08	3.45E-06
1.30E-08	2.52E-07
1.40E-08	-1.01E-05
1.50E-08	-9.35E-06
1.60E-08	-1.41E-05
1.70E-08	-1.41E-05
1.80E-08	-1.33E-05
1.90E-08	-1.57E-05
2.00E-08	-1.89E-05
2.10E-08	-1.41E-05
2.20E-08	-2.13E-05
2.30E-08	-1.49E-05
2.40E-08	-1.89E-05
2.50E-08	-1.73E-05
2.60E-08	-1.25E-05
2.70E-08	-1.41E-05
2.80E-08	-1.49E-05
2.90E-08	-1.65E-05
3.00E-08	-2.95E-06
3.10E-08	-8.55E-06
3.20E-08	-5.48E-07
3.30E-08	-8.55E-06
3.40E-08	-7.75E-06
3.50E-08	-2.95E-06
3.60E-08	-6.15E-06
3.70E-08	-1.01E-05
3.80E-08	-1.01E-05
3.90E-08	-1.09E-05
4.00E-08	-8.55E-06
4.10E-08	-7.75E-06
4.20E-08	-1.33E-05
4.30E-08	-4.55E-06
4.40E-08	-9.35E-06
4.50E-08	-1.09E-05
4.60E-08	-1.81E-05
4.70E-08	-9.35E-06
4.80E-08	-1.57E-05
4.90E-08	-1.97E-05

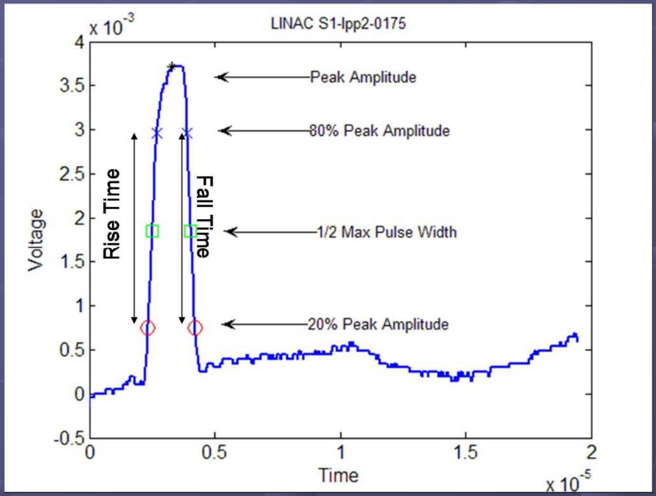




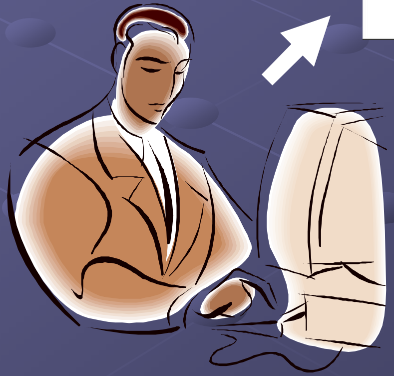
Duplication of effort

Shell script

Matlab



Perl script



C++/C function

Initial design requirements

1. Usable
2. Portable
3. Extensible
4. Flexible

Why Python instead of .....

- Matlab-like syntax
- Rapid user adoption
- Extensible via user scripts and modules
- Object oriented language

	MATLAB	Python
2X3 matrix	[1 2 3; 4 5 6]	Mat([1 2 3],[4 5 6])
Acess elements	A(2,5)	A[1,4]
Matrix inverse	Inv(A)	Linalg.inv(A)
Matrix multiply	A*B	A*B

Parameter estimation example

```
Calibrator
      DataSource model
      DataSource experiment
      Metric     comparator
```

```
From calibrator import *
C=Calibrator()


c.set_experiment('expdata.dat')
c.set_model('calore_output.dat')
c.set_metric('discrete_norm', [2])
Model_fit = c.compute_metric()
```

Related Tools

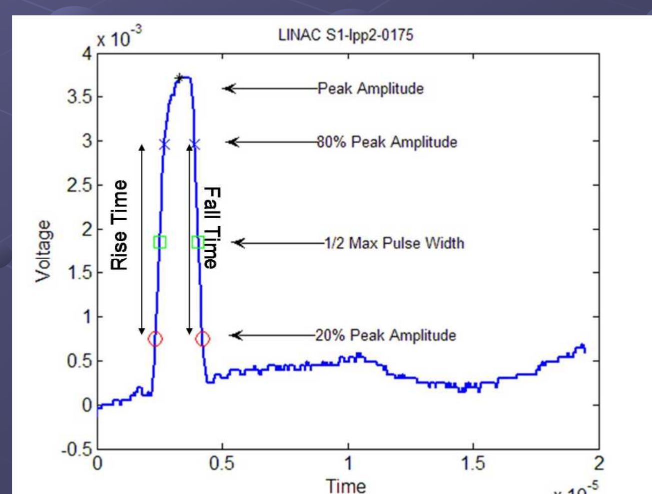
- EXACT - framework for analyzing computer experiments (Hart, Berry, Heaphy, Phillips, 2007)
  - Archiving
  - Data processing
  - Statistical analysis tool

- DAKTOOLS – framework to manage DAKOTA workflows (Tipton and Brandon, 2004)
  - Job submission
  - Analysis and management results

- For more info on DAKOTA, check out Brian Adams talk Wednesday 1:50pm in C-Room 309

A data extraction toolkit provides similar core functionality

Data
Extraction
Toolkit



Concluding remarks

- We have developed a prototype of a data extraction toolkit.
- The toolkit shows promise in reducing time spent by designers and analysts in developing customized tools.
- Potential stakeholders include mechanical and electrical engineers, and informatics analysts

Questions?

Send email to

pwillia@sandia.gov

For more info on DAKOTA, check

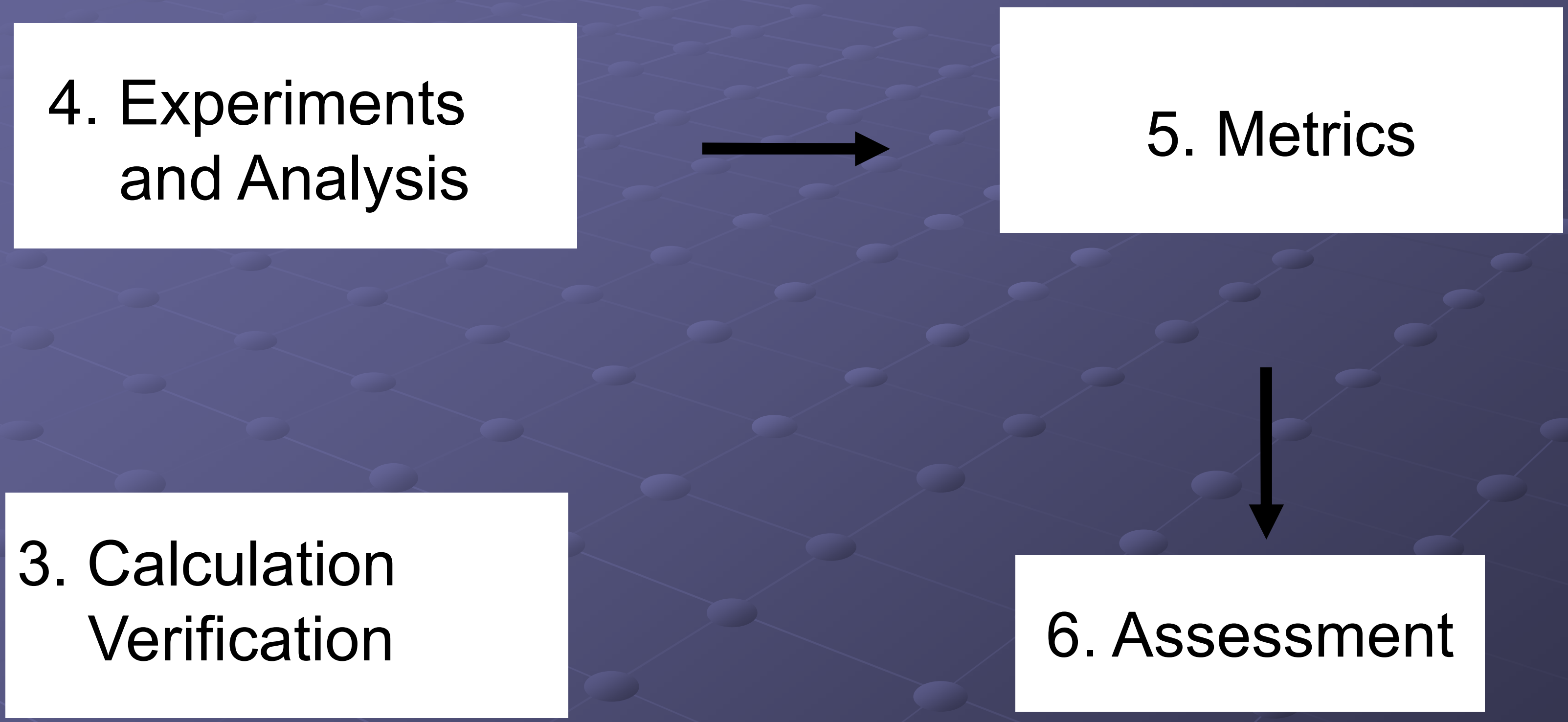
out Brian Adams' talk

Wednesday 1:50pm in C-Room 309

Existing capabilities include

- Parsers
- Extractors
- Metrics

Validation experiments



Motivation (Change for pymedics)

- Facilitate algorithm development and comparison
- Define common interface to different methods
- Provide an extensible infrastructure for new algorithms or problems
- Provide reusable optimization components and codes