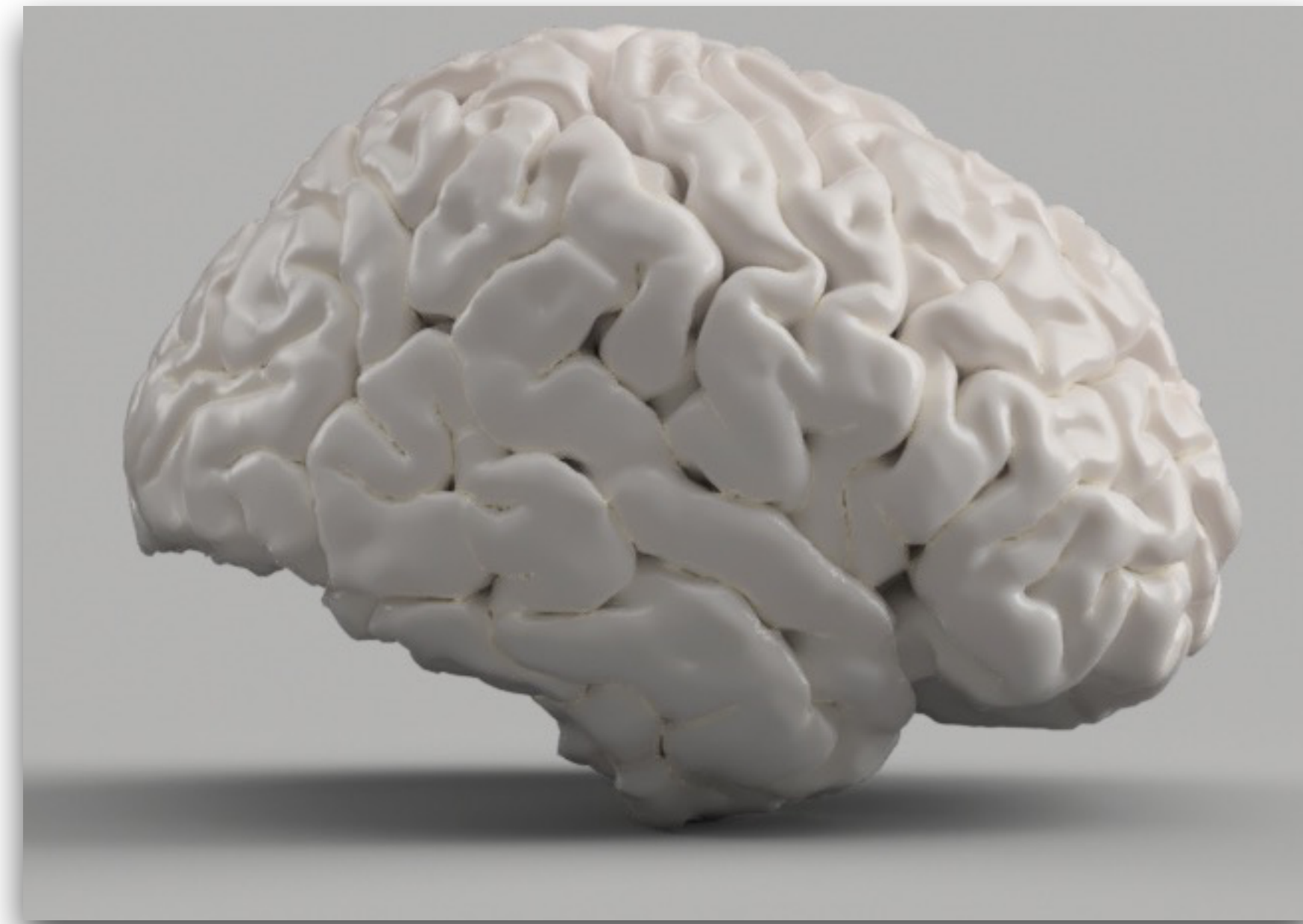# Welcome to CogSci 109

Modeling and Data Analysis
Summer Session 1, 2023
C. Alex Simpkins Jr., Ph.D.
RDPRobotics LLC | Dept. of CogSci, UCSD

Instructional Team

- **Instructor**
  - C. Alex Simpkins PhD
    - rdprobotics@gmail.com, csimpkinsjr@ucsd.edu
- **TA**
  - Sagarika Sardesai
    - ssardesai@ucsd.edu

C. Alex Simpkins Ph.D.

Plan for today

- Introductions

- About your instructor and what I'll try to share

- Motivation for the course

- Defining Modeling and Data Analysis

- The structure and mechanics of the course, assignments, project, etc

- Expectations, how to succeed in the course

Who am I?

- C. Alex Simpkins Jr. Ph.D.

- BS/BS/MS/PhD UCSD Psyc, AMES, MAE, MAE, 2 postdocs UCSD Cogs and UW CSE

- Taught as TA ~20 times as a student, taught COGS109 as a grad student, taught at SDSU for a year in ME in Design, came back to UCSD Winter quarter - COGS100, 108 and 138

- Been involved in teaching for over 30 years teaching martial arts

- PhD in Control Theory (and Design) involving a great deal of modeling/data analysis

- Machine learning

- Robotics
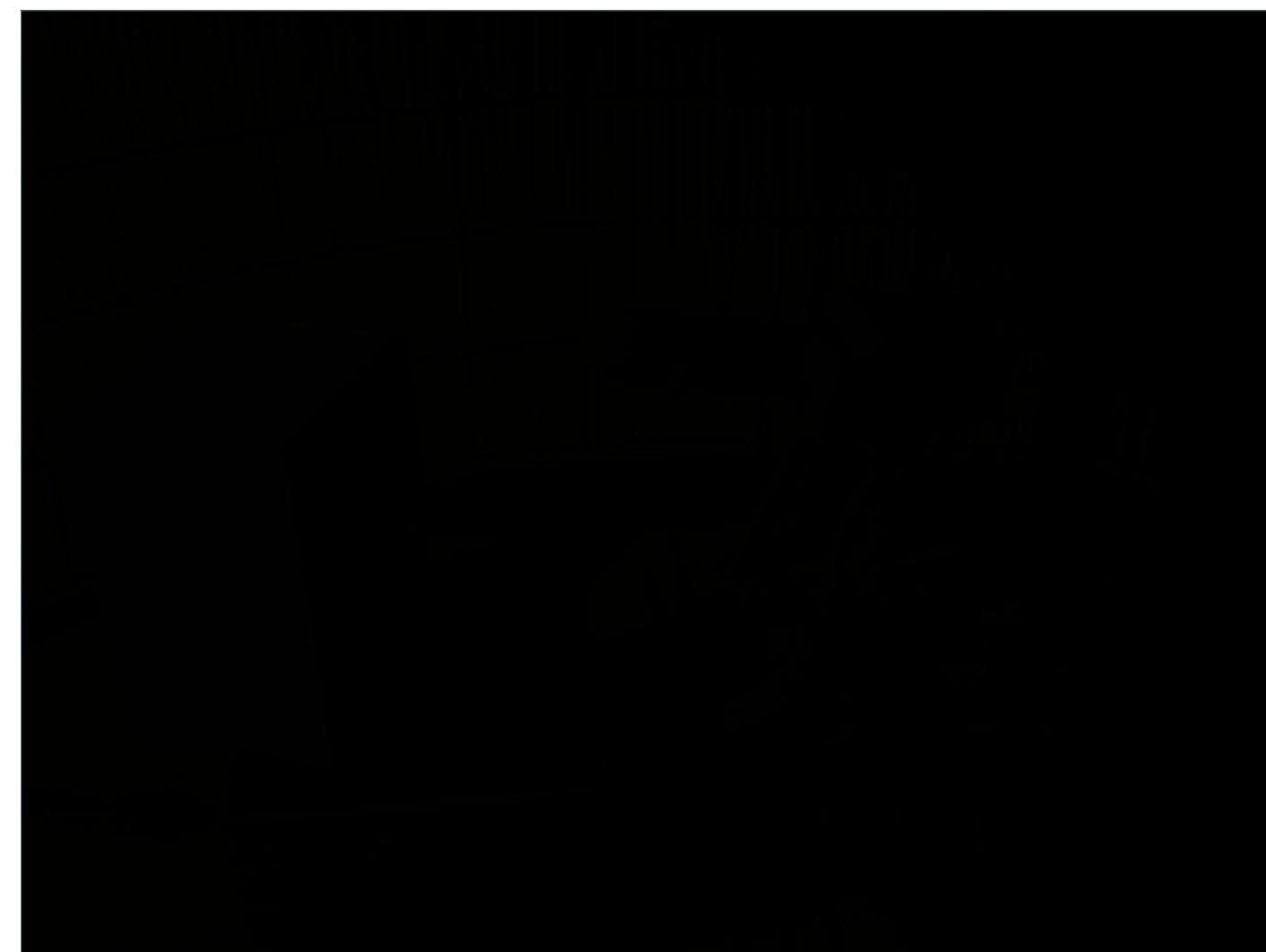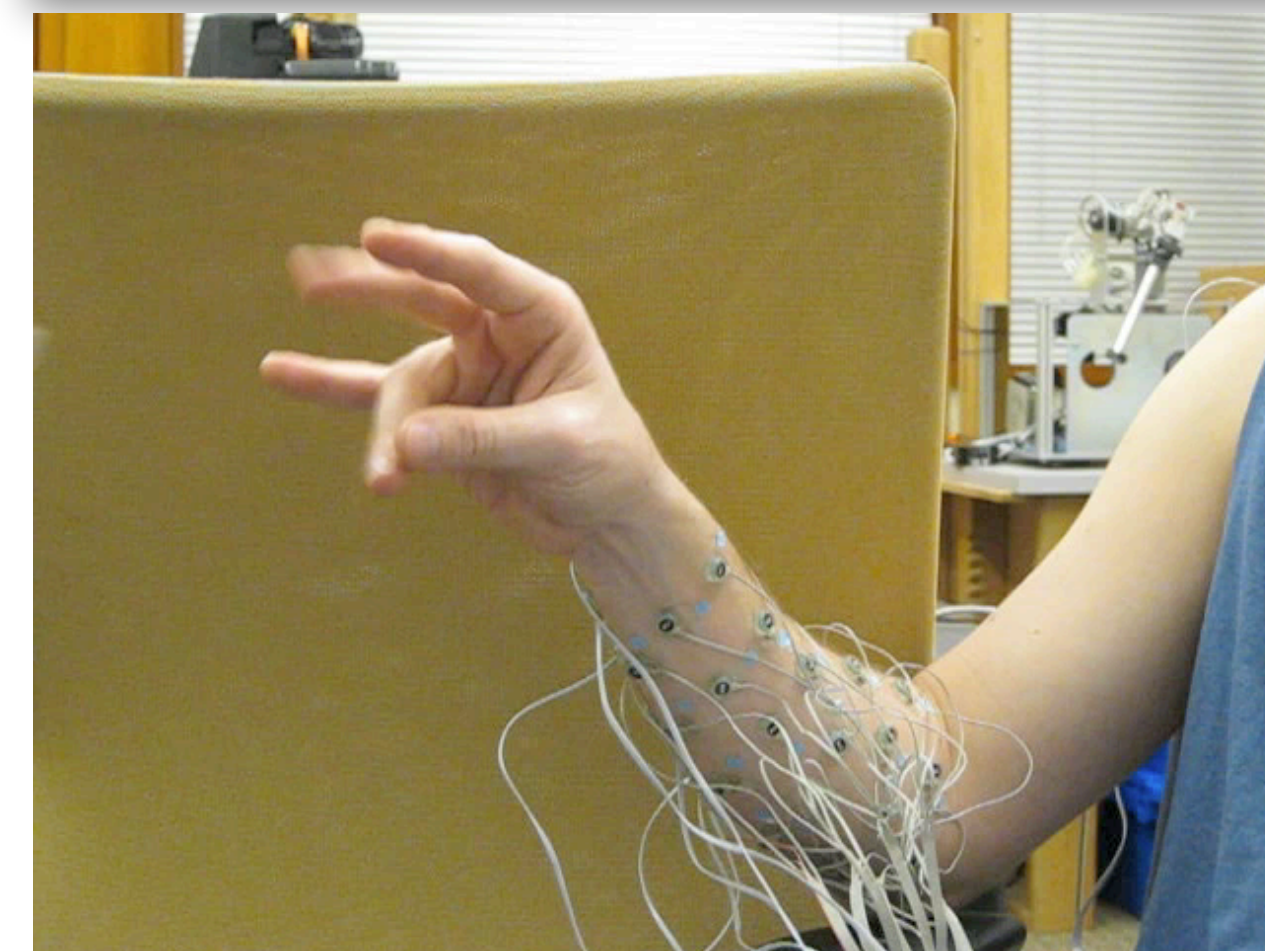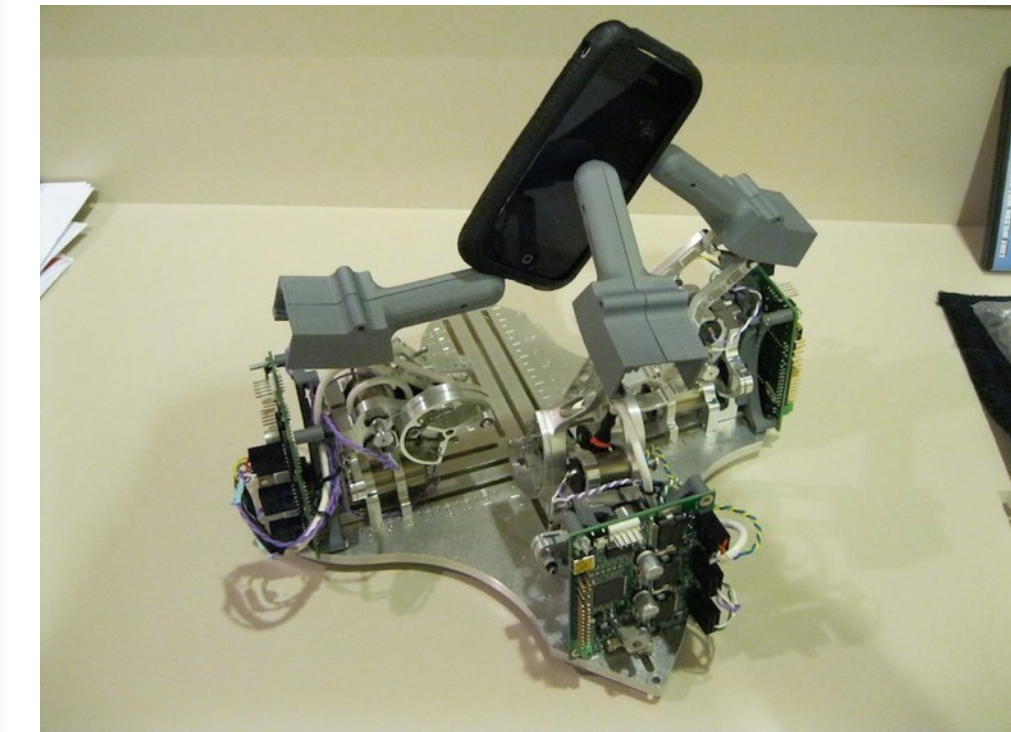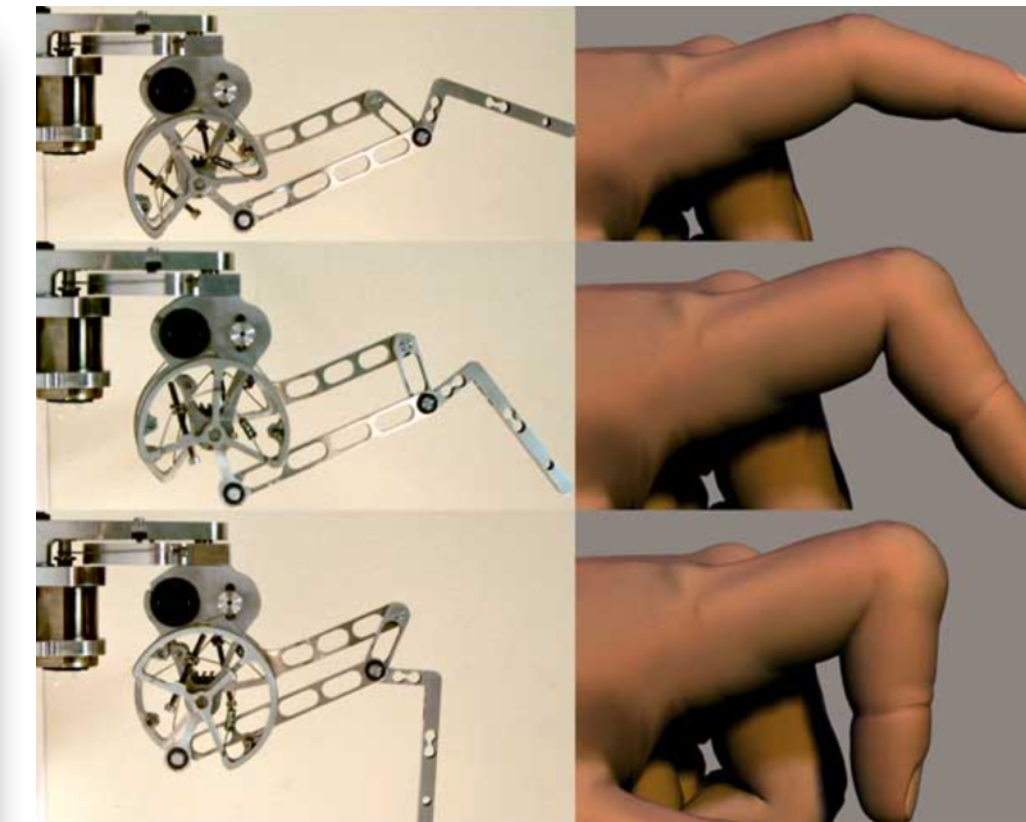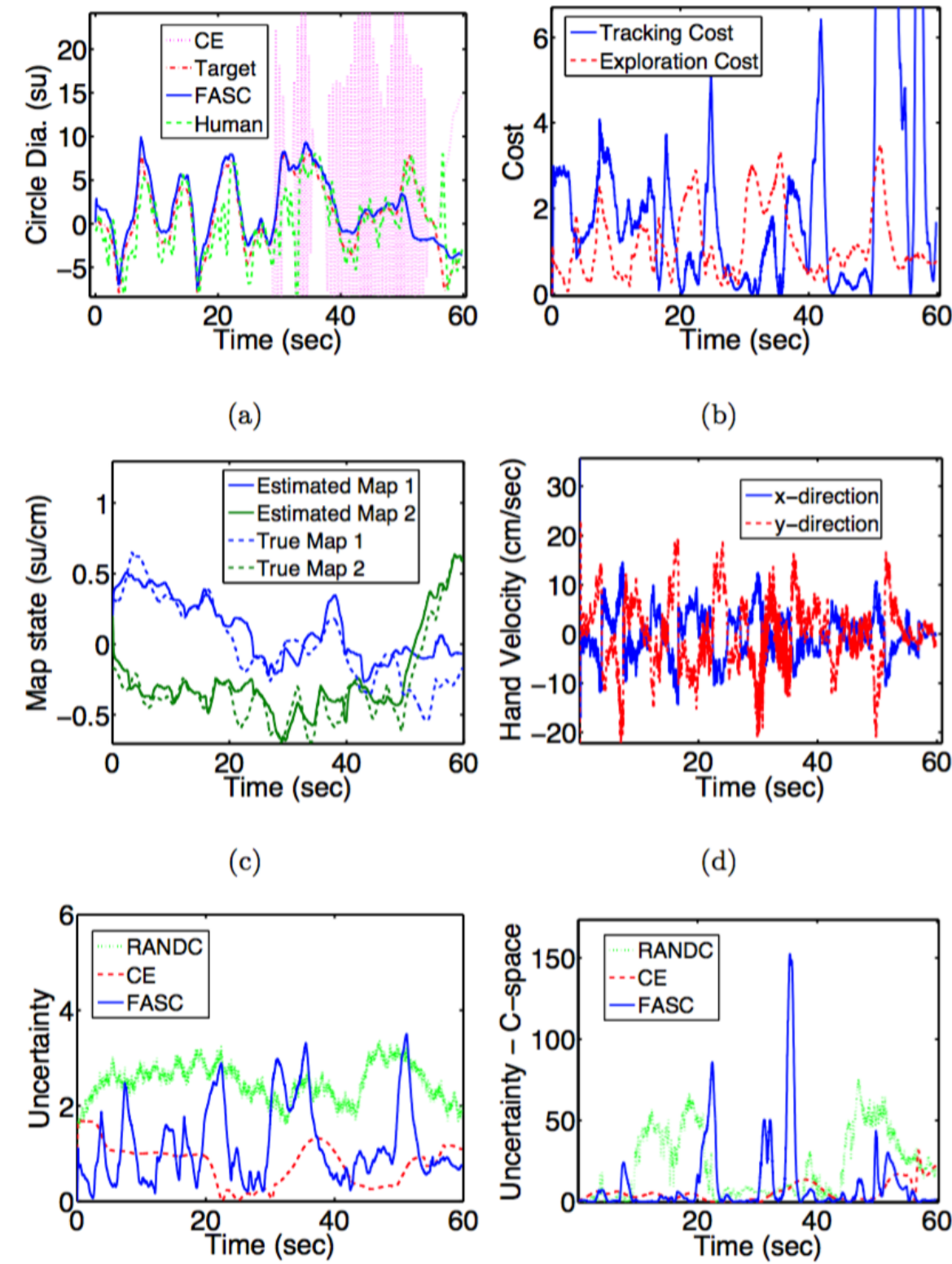
Who am I?

- Many industrial projects - robotics, AI, drones, sensor systems
- Consulting
- Entrepreneurial startup work
- Work with larger companies
- Research - COVID, robotics, AI, control
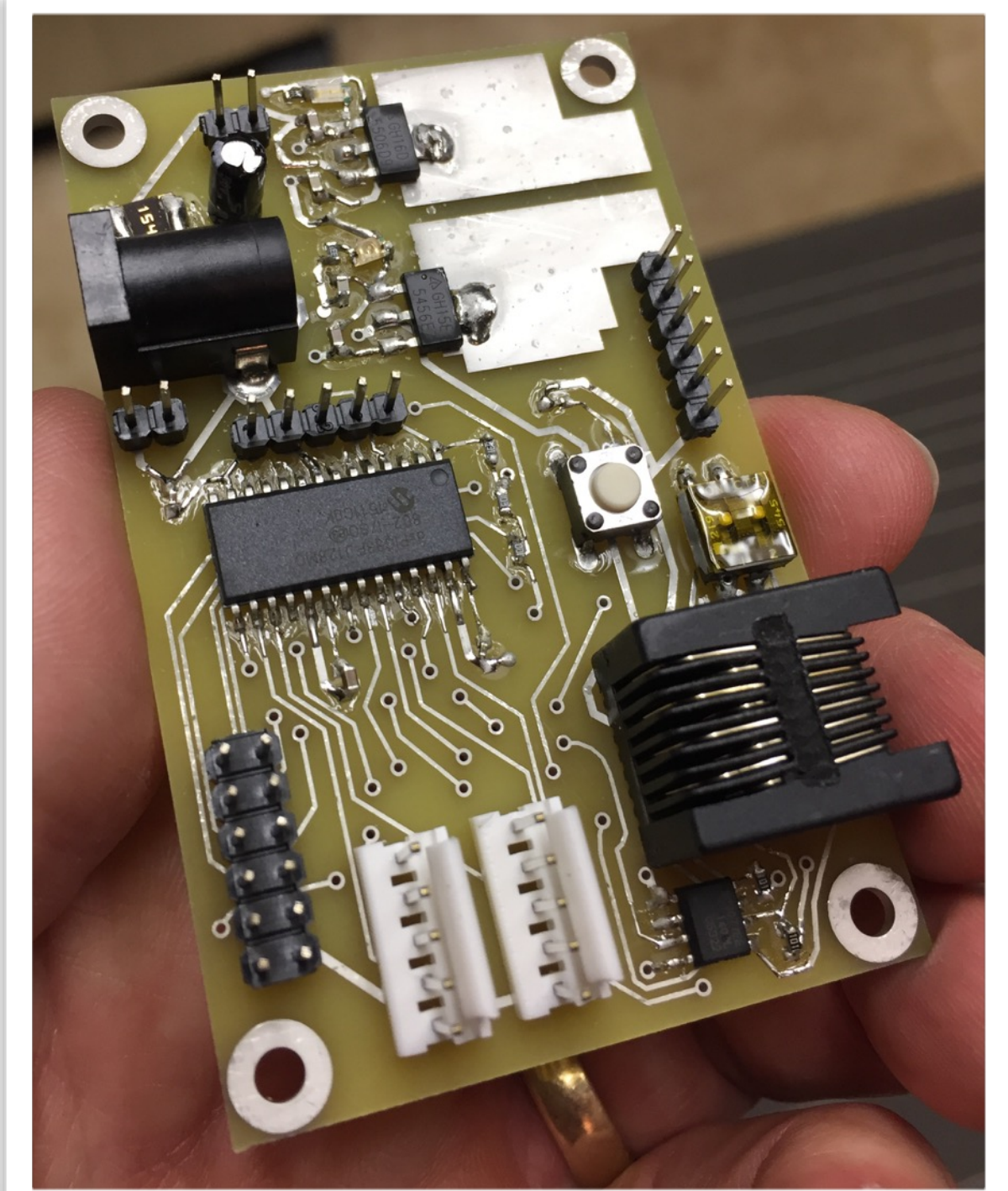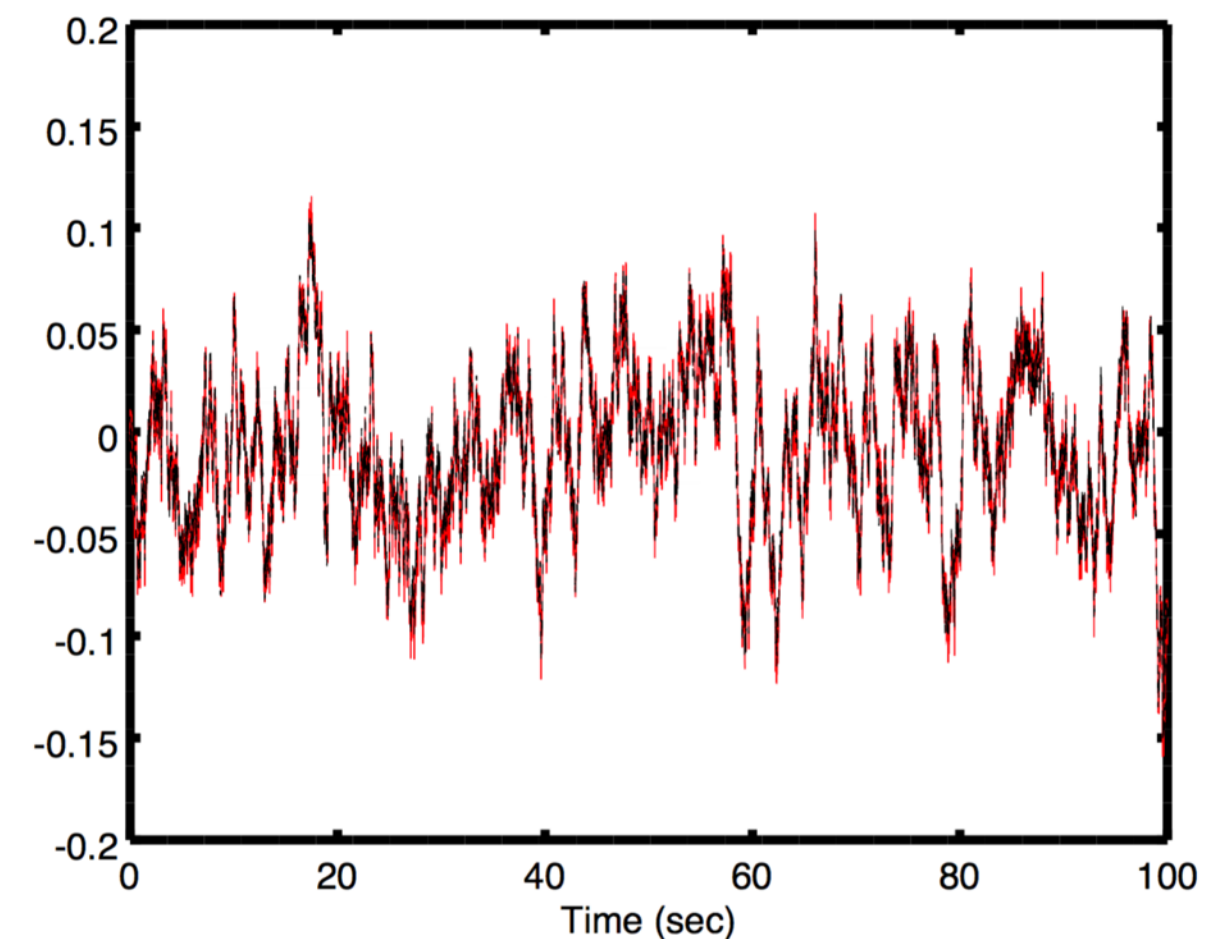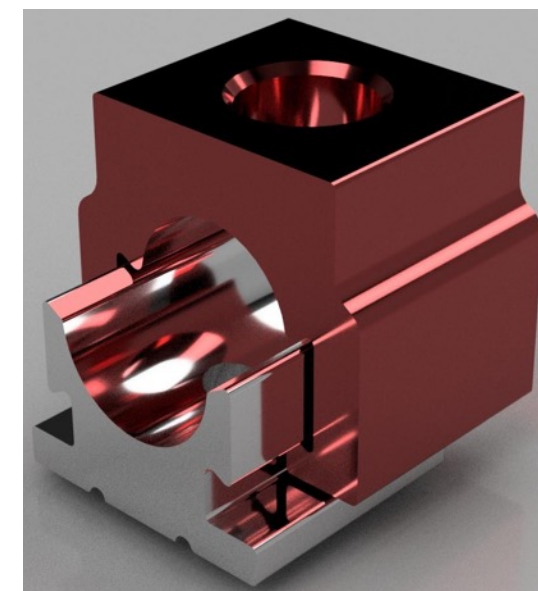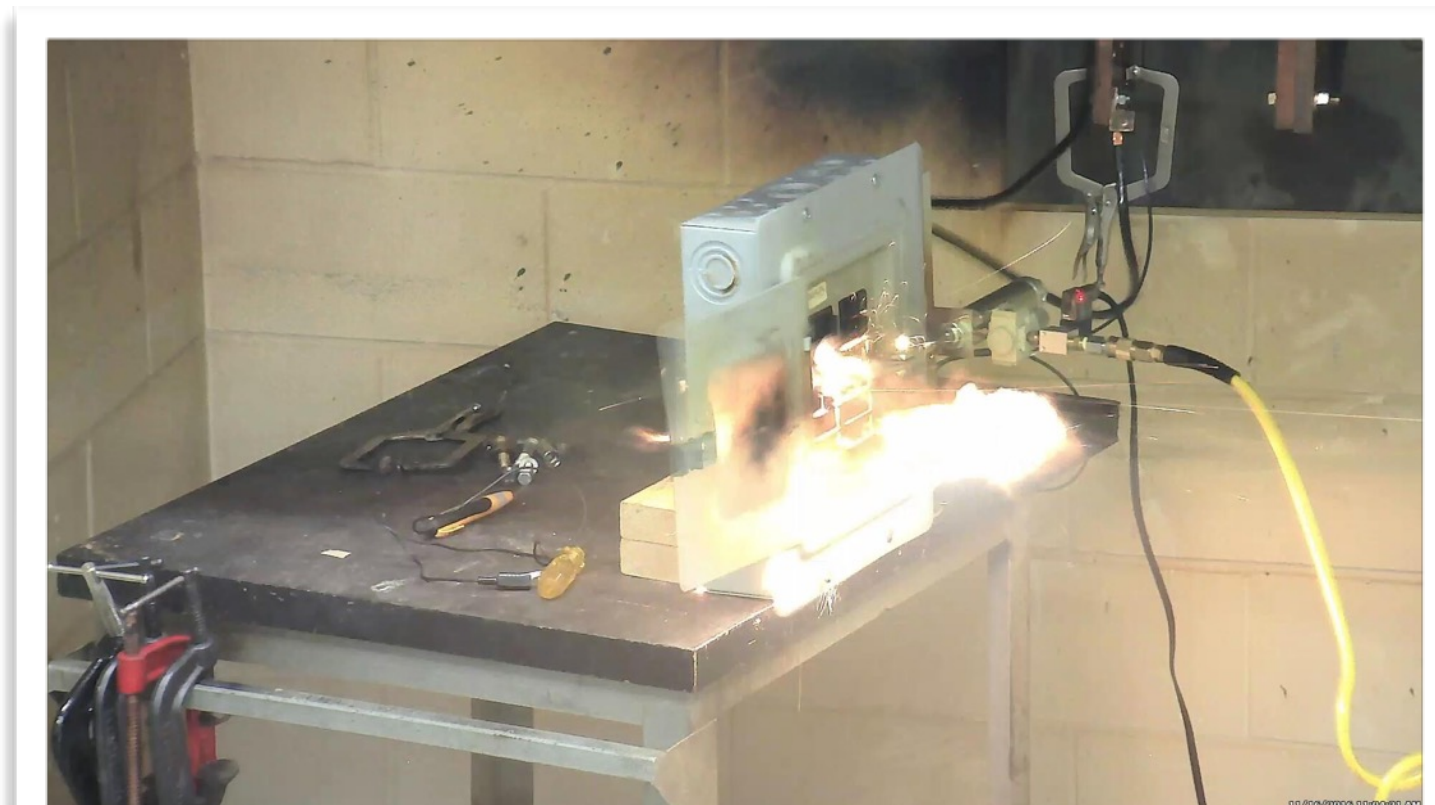
What to expect in this class

- What is my (our) role?

  - I am (we are) here to help you to learn and succeed, to open the door

  - NOT here to weed anybody out

  - NOT here to compete with you

  - Mutual respect

What to expect in this class II

- What is your role?
  - Learn! Open your mind
  - Put in the effort - you must walk through the door
    - Watch/attend lectures, do the readings, complete assignments and tests, and think about it all
  - Treat each other well, help each other to succeed (but do your own work of course)

Our expectations of you

- This course does not assume that you lived every day programming since you took COGS18

- Assuming that you took the intro to programming or are taking concurrently, and so have some sense of what programming is, but do not necessarily remember all details

- We will review and define mathematical and programming concepts as we go

- It will require an open mind on your part and effort

Motivation

- In Cognitive Science, we consider questions regarding the nature of cognition

  - In our department we focus on studying these questions through the brain, behavior and computation

  - In other fields of science and engineering, one considers other questions, but the techniques are mostly the same for data, analysis and modeling

  - Complicated! In order to determine across individuals a reasonable response to these questions, one must be able to describe patterns which are beyond the subjectivity and human limitations of the individual

  - Therefore, in CogSci (and other fields) we use tools of modeling and data analysis
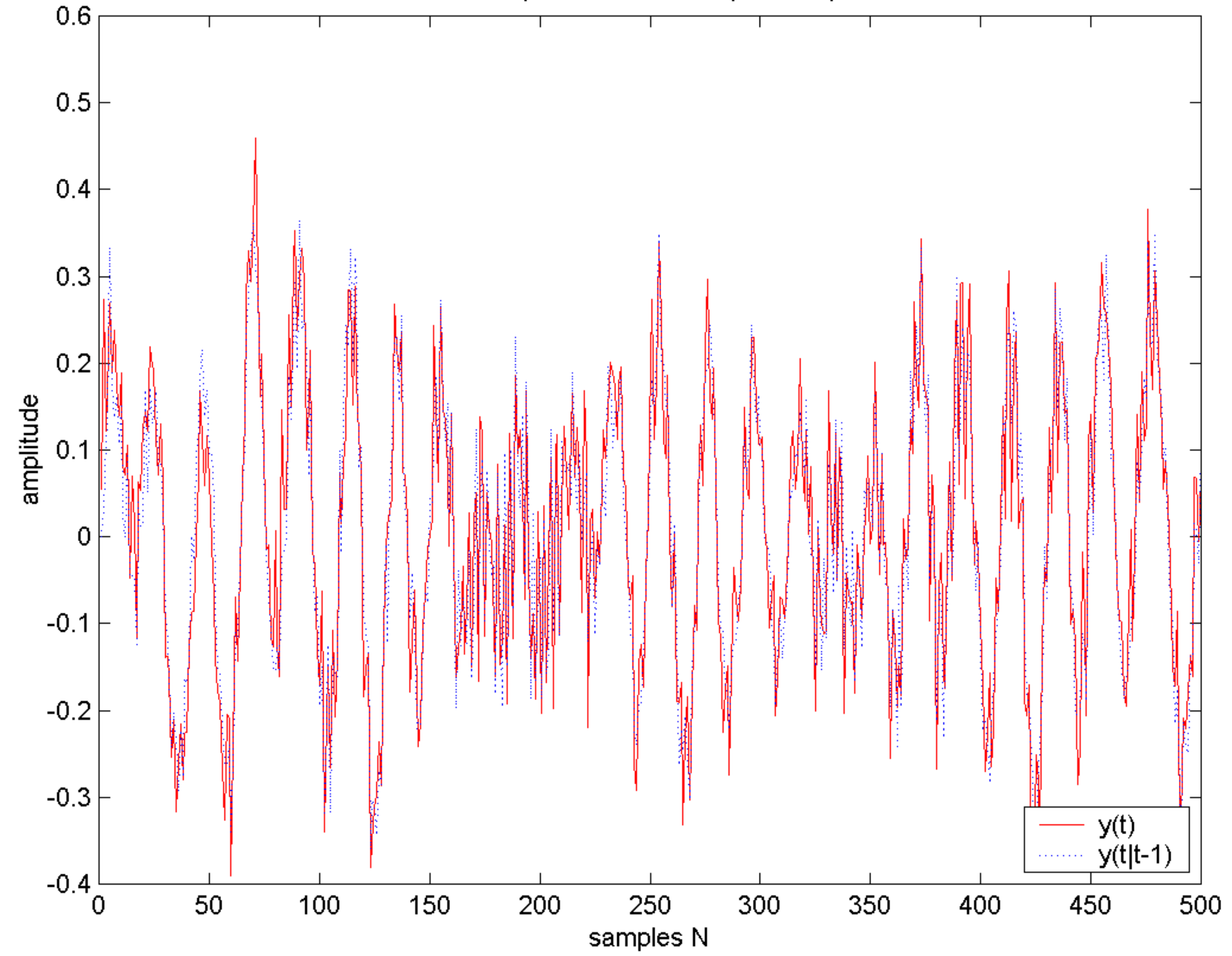
Bode - actual system vs. 4th order ARX
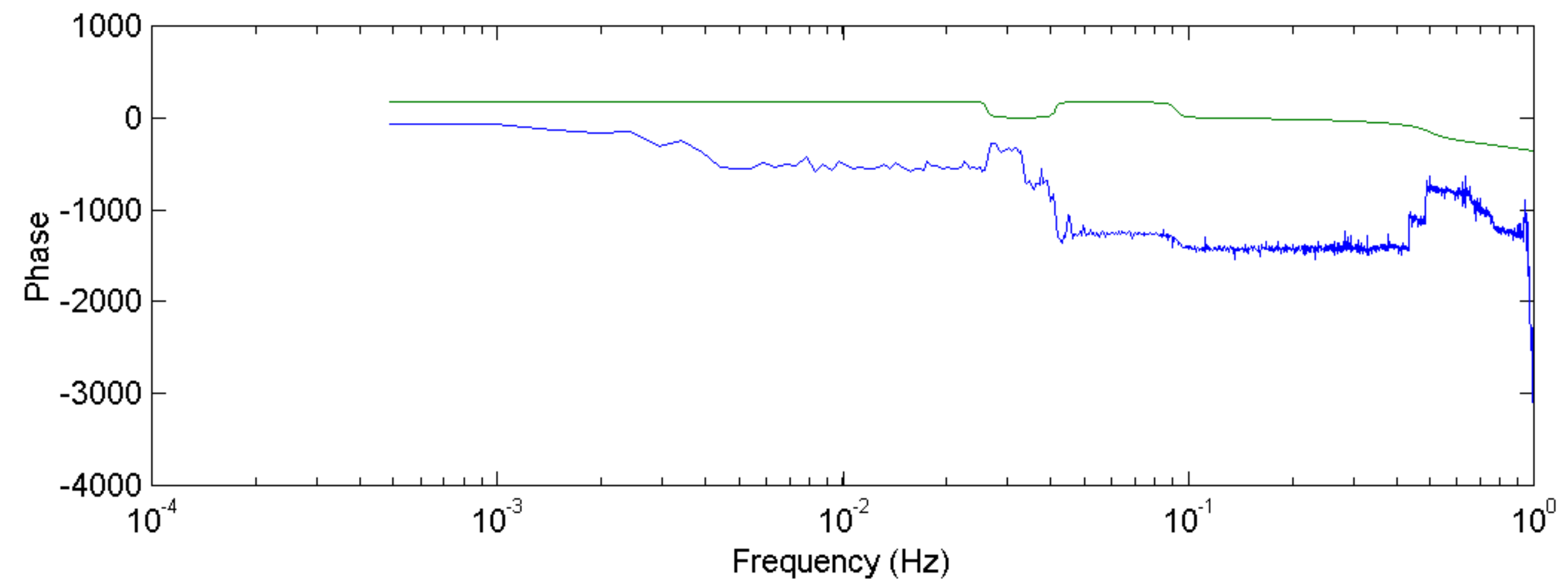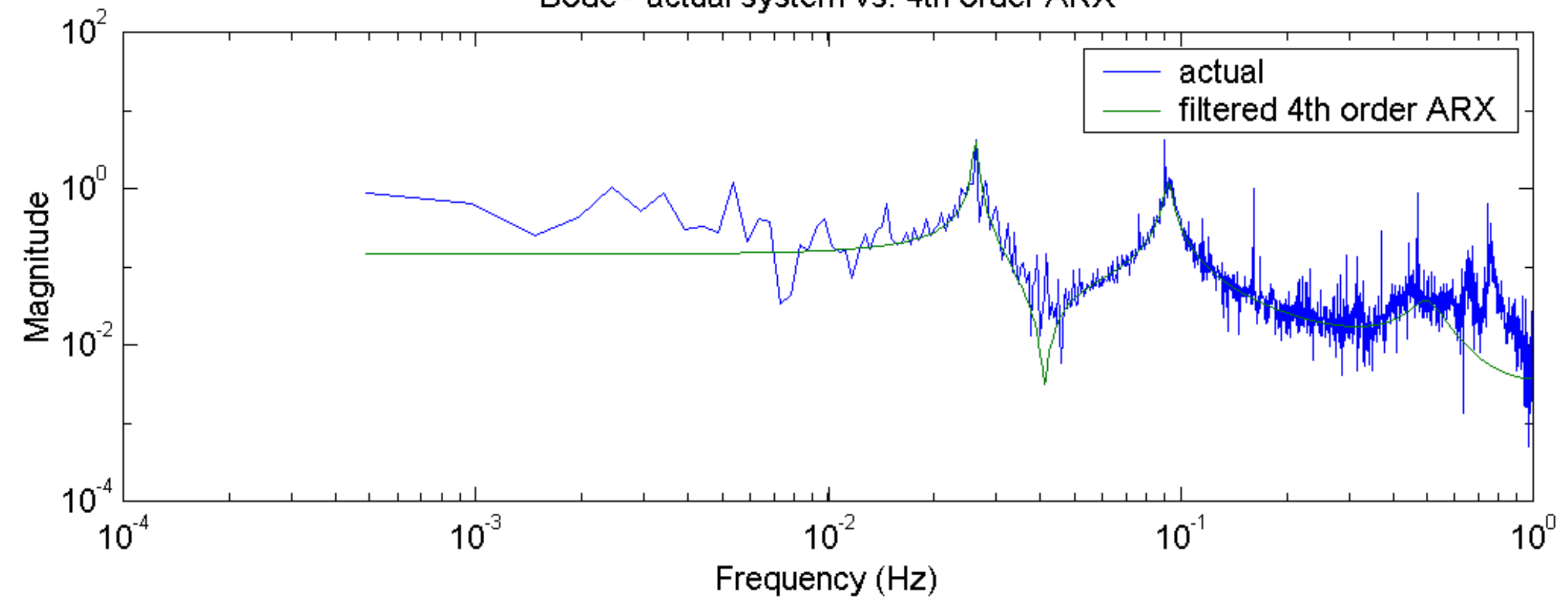
What is '*Data*?'

- "Facts and statistics collected together for reference or analysis" (Webster's Dictionary)

- Can be recorded from any form, situation, or field of study

- "Anything that exists, exists in some quantity, anything that exists in some quantity can be measured…" (Thorndike)

What is '*Analysis*?'

- "detailed examination of the elements or structure of something, typically as a basis for discussion or interpretation" (Webster's Dictionary)

- "the part of mathematics concerned with the theory of functions and the use of limits, continuity, and the operations of calculus"(Webster's Dictionary)

What is '*Modeling*?'

- "a system or thing used as an example to follow or imitate" (Webster's Dictionary)

- "a simplified description, esp. a mathematical one, of a system or process, to assist calculations and predictions" (Webster's Dictionary)

- **Predictions** - once you have a good model, you can use it to help answer fundamental questions by creating predictions of the specific cognition, behavior, etc you are studying, and compare predictions to reality

Putting it together

- In modeling and data analysis we take a system that we want to make sense of, predictions of, or decisions with, and perform computational analysis on the data available, starting with basic observations and expanding to complex models.

- We are trying to understand connections, patterns and maps in order to make statements about the system

Why this course is important for everybody

- No matter what you do in life, being able to take in information, organize it, do something with it, and then communicate with others is key

- These specific techniques are very basic to scientific research

- It is a language to speak and understand

Survey (link also on canvas and web page)

COGS109 student survey (SS1 2023)

This survey is used to help me get to know you better! Thanks for your participation

Complete before Fri. week1 @ 11:59pm - opportunity for a little extra credit

If any data is used in class, the data will be anonymized. How you respond will not affect how you do in the class. Many are not required questions, please do not answer questions that make you uncomfortable

your email address will be recorded when you submit your form

How we'll approach learning about *and doing* Modeling and Data Analysis in COGS 109

Scheduling & Staff

Lecture: Tu, Th 11am-1:50pm in SOLIS 109
Discussion Sections: T, Th 2-2:50pm in SOLIS 109
Office Hours: TBA(Dr. Simpkins, by appt.);

Instructor			TA	
C. Alex Simpkins Ph.D.	rdprobotics@gmail.com		Sagarika Sardesai	ssardesai@ucsd.edu

Five main parts to the course

Week #	Topics
Week 1	Introduction and definitions, datahub, python, version control/git/github
Week 2	(Python/Jupyter II) Data manipulation and processing
Week 3	Extracting basic information from data and visualizing that info
Week 4	Modeling the data and evaluating models, data fits
Week 5	Presenting and communicating results

Programming prerequisite

- MAE 8 - MATLAB
- CSE 8A or 11 - Python/Java
- COGS 18 - Python
- DSC 10 - Python

Bottom line: we will assume programming knowledge.

Python will be used for all labs/projects/assignments.

No programming experience (or you forget it all)?

- *Preferred option*

  - Take a programming course first

  - COGS 18 : Introduction to Python

- *Can't wait?*

  - Use online sites like *codecademy.com* or *LearnPython.org*

  - *Python Data Science Handbook*

- *We will work with you to assist you - this is not as heavy a programming course as COGS108, and we will be leading you through what you need to know*

Course links

Web page	http://casimpkinsjr.radiantdolphinpress.com/pages/cogs109_ss1_23/index.html	**Central hub**: lecture/section materials, readings, links and jumping off point to the other resources
datahub	https://datahub.ucsd.edu	**assignment** submission
Piazza	https://piazza.com/ucsd/summer2023/cogs109_s123_a00/home (course code on canvas home page)	**questions**, discussion, and regrade requests
Canvas	https://canvas.ucsd.edu/courses/47870	grades, lecture videos, quizzes, mirrors of lectures
Course feedback	Submit via google form	if I ever offend you, use an example you hate, or to provide general feedback

Attendance is not required, but live interaction is strongly encouraged

- <u>All lectures</u> will be recorded (available by 2PM every TTh; Canvas Media Gallery)

- The technical discussion section each TTh will be recorded

- Zoom for live questions

- Office hours, piazza, email

Assignment deadline timing

- When possible (which is most of the time), we will give you a week for an assignment

- Summer session is a little compressed, so for some assignments we recommend you complete a week after an assignment

  - Soft deadline (week after assignment)

  - Hard deadline (Fridays) - for simplicity

Weekly Lecture Quizzes:

- (4-5) weekly quizzes (first one due Friday of Week 2)

- Goal: to help you keep on top of the material covered in lecture

- Why?: experience + student feedback

- How:

  - Taken on Canvas

  - Multiple Attempts (3)

  - ~10 Questions

  - Timed : 30 minutes

  - Posted by Friday @ 11:59 PM (after each week of lecture); due the following Friday

  - Meant to test concepts from previous week's lecture

Lecture quizzes will be due on Fridays by 11:59 PM.
If we get through all 5, one will be extra credit

(4) Assignments

Assignments are completed individually and graded programmatically.

- These are meant to get you practice programming around the topics covered in class.
- The first two are much simpler than the following two and should take less time.
- You will have to look some things up on your own. This is by design, and a skill to develop!
- Instructions must be followed to receive credit.
- You'll have the opportunity to practice in discussion section.

Assignments will be due on Fridays by 11:59 PM.

75% credit if submitted w/n 72h after deadline, subject to instructor/TA judgment

Assignment Submission @ Datahub: https://datahub.ucsd.edu

DATA SCIENCE / MACHINE LEARNING PLATFORM

UC San Diego

Information Technology Services - Educational Technology Services   Help Options ▼

jupyterhub

Log In
Registered Users
"username@ucsd.edu"

UC San Diego Jupyterhub (Data Science) Platform

Before: log onto datahub & have a working installation of Jupyter on your computer

Discussion Section

- Goals:
  - More opportunity for individual contact
  - Help with technical aspects of the course
  - Assignment & project help
  - Labs submitted by Fri @ 11:59 PM (2%/lab; 1 may be EC if needed)
- Can I switch sections? Nope there's only one!
  - Do lab exercises on your own if you are comfortable with material
  - Questions via piazza if can't attend
  - The section is always recorded

Discussion Sections will start today!

Logistics of the course

Assignment type	Grade weight	Form
Assignments (4)	32%	Datahub
Discussion workbooks (8)	16%	Datahub
Quizzes (canvas, 4-5)	20%	Canvas, remote
Project (1)	32%	github
Extra credit	~10%	remote, SONA
Total:	100% +10%bonus	—

Logistics of the course

- Course page
  - http://casimpkinsjr.radiantdolphinpress.com/pages/cogs109_ss1_23/index.html

- Grading (Fill the bucket ~ 1000pts)

- Each week, due Fridays @11:59pm, turn in to https://datahub.ucsd.edu

- Final project (group)

- Bonus (10% total)
  - SONA (2%), extra quiz (2%), surveys/class participation (6%)

Assignment type	Grade weight	Form
Assignments (4)	32%	Datahub
Discussion workbooks (8)	16%	Datahub
Quizzes (canvas, 4-5)	20%	Canvas, remote
Project (1)	32%	github
Extra credit	~10%	remote, SONA
Total:	100% +10%bonus	–

When is the first thing due?

- **This Friday**

  - Pre-course survey - posted today/tonight after class, please complete by Friday for some extra credit

  - Group survey - posted today

- **Next friday at 11:59pm**

  - Quiz 1 (Recommended by Wed - will not take a long time)

  - D1 (Recommended soft deadline Wed)

  - A1

  - Paper review (group - actual review fairly short)

Course Confusion and getting questions answered

- If something in lecture, a section workbook, or an assignment is unclear:

  - *Ask in class*

  - *Ask during section*

  - *Post on piazza*

  - *Ask a classmate*

  - *Come to office hours*

Canvas messages are less ideal - email or piazza first.

We are going to focus on piazza first then email.

Clarification on communications

- I do my very best to be approachable as we all will
- I have many students this quarter (not to mention two small children and a small company, research, and more going on)
- If I am every slow or brief in a response, it's probably that I have many communications to respond to, it is never to be taken as anything else

CLASS CONDUCT

- In all interactions in this class, you are expected to be respectful. This includes following the UC San Diego principles of community.

- This class will be a welcoming, inclusive, and harassment-free experience for everyone, regardless of gender, gender identity and expression, age, sexual orientation, disability, physical appearance, body size, race, ethnicity, religion (or lack thereof), political beliefs/leanings, or technology choices

- At all times, you should be considerate and respectful. Always refrain from demeaning, discriminatory, or harassing behavior and speech, including to ANY of the teaching staff in class or online. Last of all, **take care of each other**.

- If you have a concern, please speak with Dr. Simpkins or your TA. If you are uncomfortable doing so, the OPHD and/or CARE are excellent resources on campus.

What you will be able to do by the end of the course

- Use Python, Jupyter, Datahub, GitHub, and other tools
- Recognize, Categorize, and Manipulate data of various forms
- Load
- Sort
- Filter
- Rearrange/size, etc

What you will be able to do by the end of the course (continued)

- Extract basic information from that data

- Standard statistics (mean, median, mode, standard deviation, correlation, variance, covariance, etc)

- Basic Hypothesis testing

- Visualize the data from multiple perspectives

- 2D plots

- 3D plots

- Charts

- Color plots/contour plots

- Basic volume visualization

What you will be able to do by the end of the course (continued)

- Create models from the data which approximate the behavior of the system

- Using linear and nonlinear methods of fitting models

- Least squares (regression)

- Basic function minimization and optimization (for example using scipy.optimize)

- Gradient descent

- Multiple criteria for fit properties (goodness of fit, etc)

- Perform basic simulations

What you will be able to do by the end of the course (continued)

- Communicate results effectively
- Communicate with others
- Read and understand the literature, speak/understand the language
- Know how to look for more information
- Expand your knowledge
- Where to go from here
- Pitfalls to avoid

Logistics (continued)

- **Labs**
  - Discussions after lecture Tu-Th 2-2:50pm same room (SOLIS 109)
  - Review material from the week, hands on experience, Q&A, workbooks/discussion labs, ***recorded***
- **Office hours locations and dates/times TBA**
  - 2 for Sagarika and by appointment
  - 2 for Dr. Simpkins and by appointment
  - we'll do a survey to see what times are best
- We're here for you! We want to help!

We want you to work together, but...

- Please don't cheat!

- That said, we need to define what's ok, what's not

  - On homeworks - you are welcome to discuss material, homeworks and so on, but you must write your own code and write-ups (no copy and pasting)

  - On tests - no collaboration, no discussion with anyone but TA's or instructor. They may give hints but not answers

- Standard UCSD academic honesty policies also apply

A final word…

- We are here to do our best to help you learn and succeed

- My goal is for you to walk out after the course having learned something that becomes part of you, helps your future

- We all want this to be a positive experience for you that you can continue to gain from over the years to come

Do you have any other logistics questions about COGS109?

I'm excited to have you all in the course.

Welcome to COGS109!

Let's take a 10min break and come back for lecture 2