Please install our GUI!

1. Go to github.com
2. Search SteveGaemi/SR-ASB
3. Read through README.md and follow instructions
4. If you have any questions, please ask us during breaks!

Wifi: ASB 2025

Password: Mayoishiring (with a Capital M)

Discovering Equations of Human Behavior

Aaron D. Likens, Ph.D

Seung Kyeom Kim, B.S

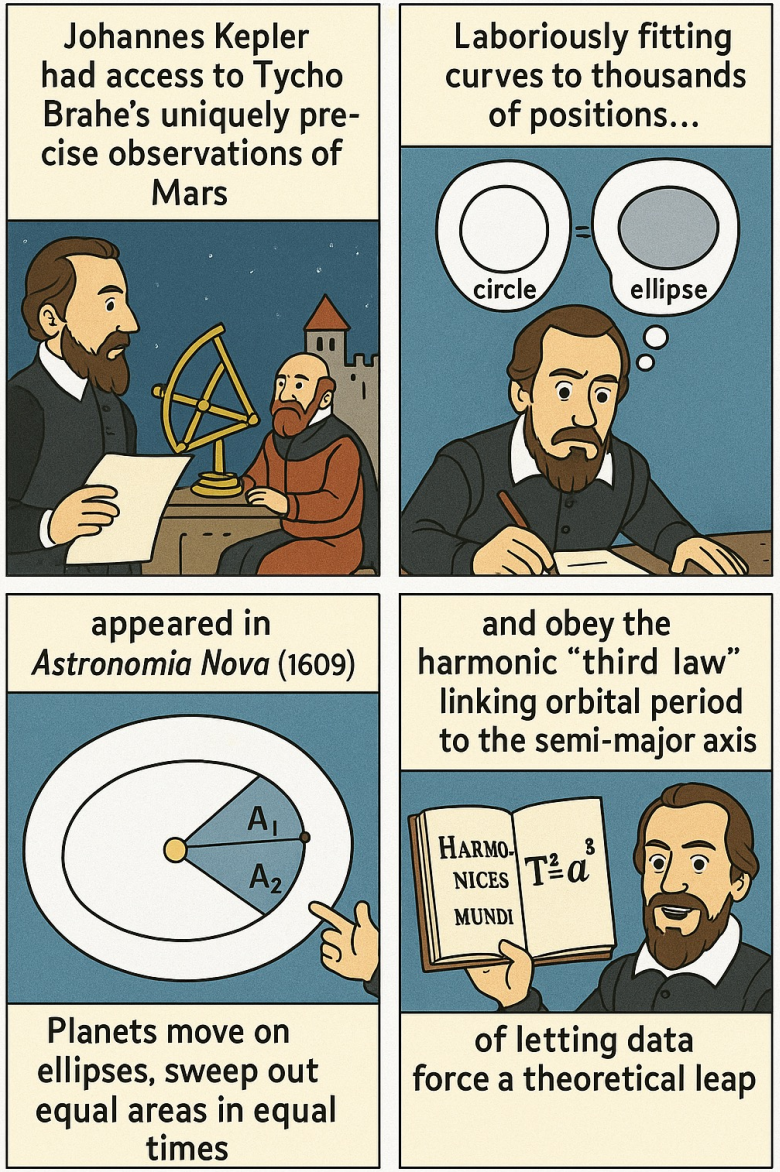
University of Nebraska Omaha

Genetic Programming

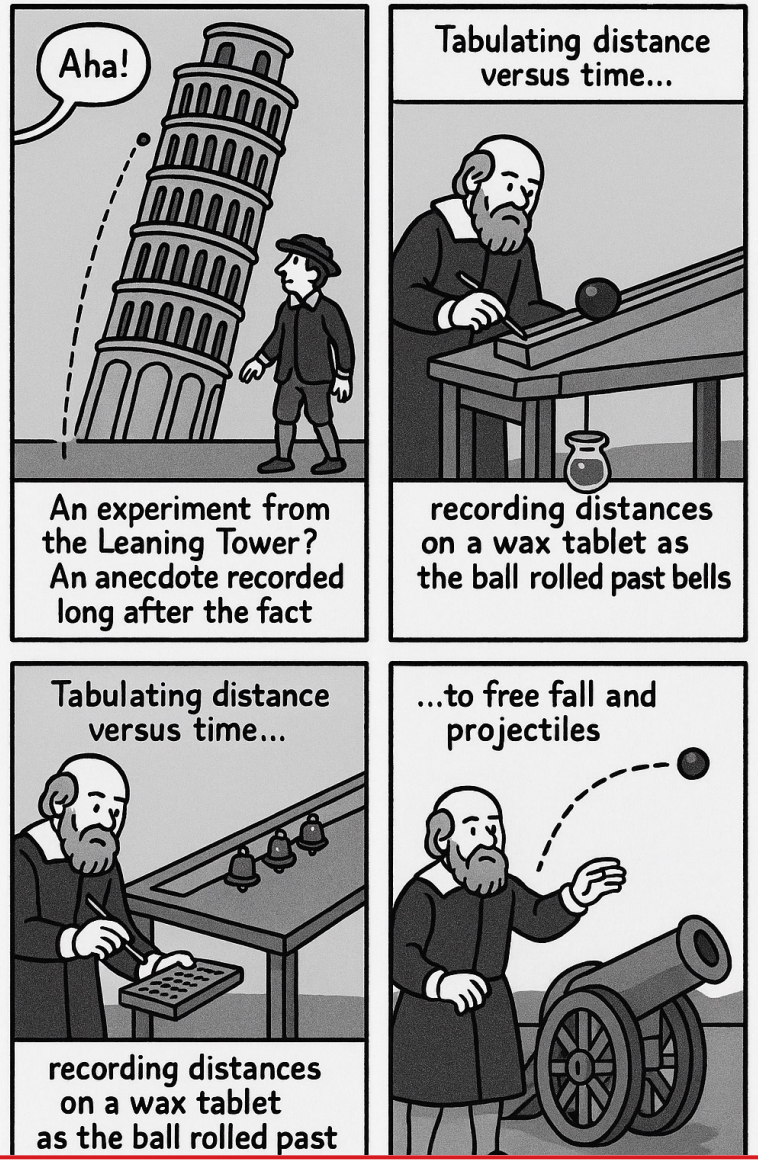
Discovering Scientific Laws from Data

Discovering Scientific Laws from Data

Discovering Scientific Laws from Data

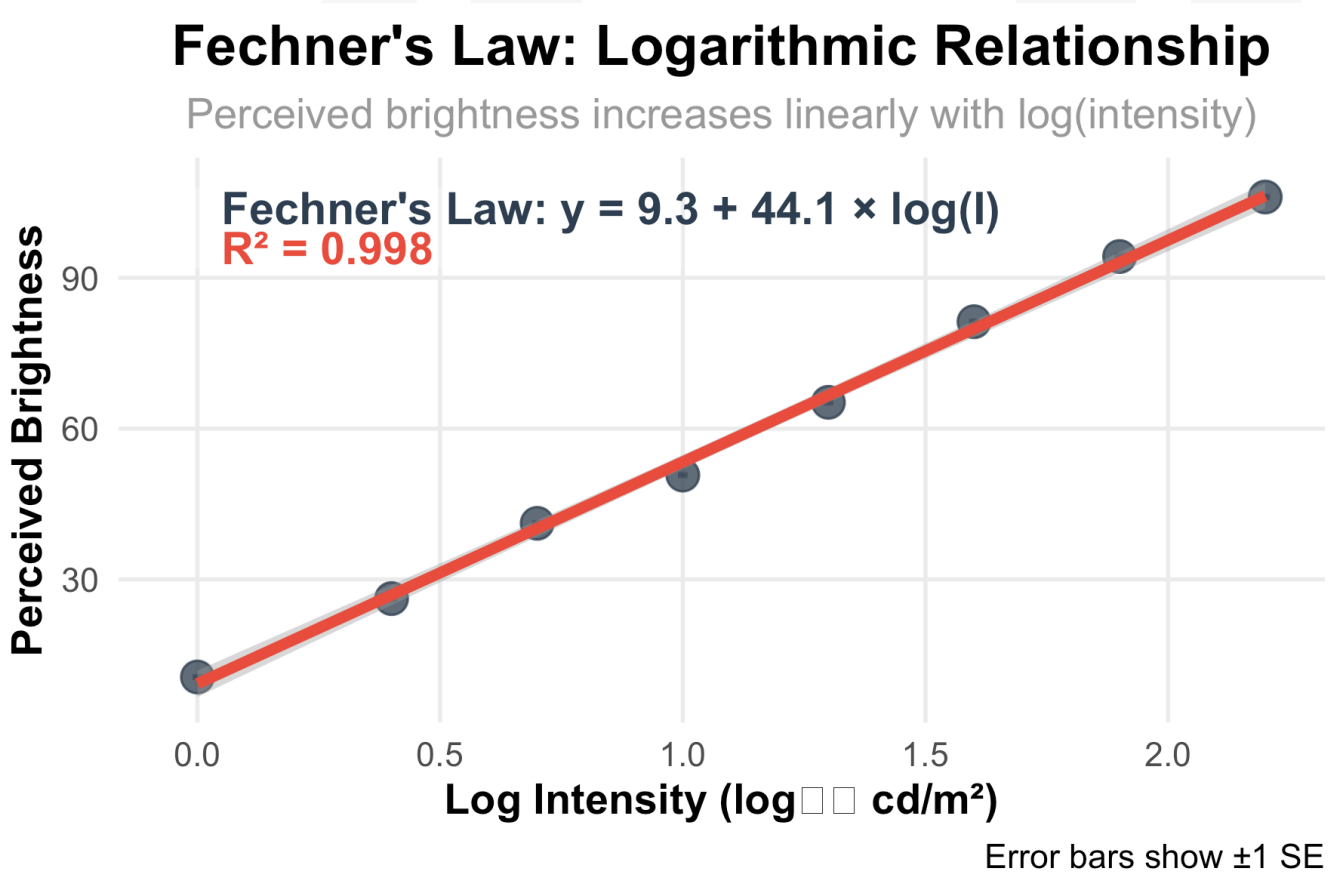


Discovering Scientific Laws from Data

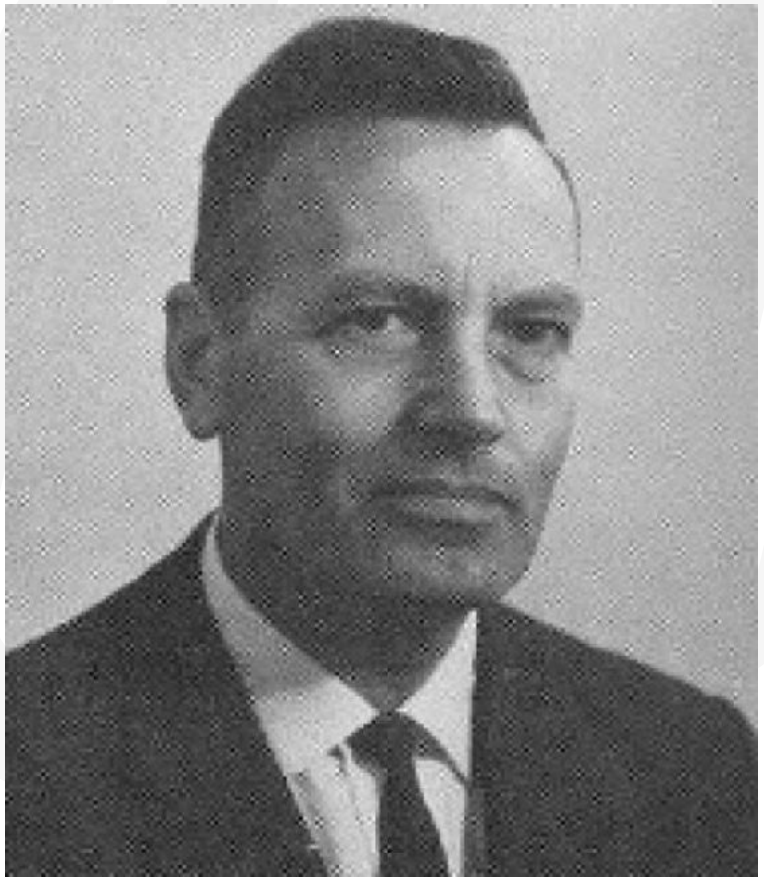


Discovering Scientific Laws from Data

Discovering Scientific Laws from Data



Discovering Scientific Laws from Data



Discovering Scientific Laws from Data

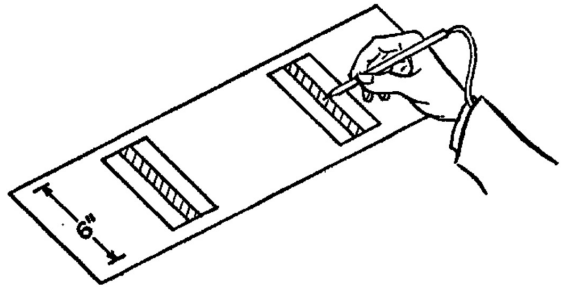


FIG. 1. Reciprocal tapping apparatus. The task was to hit the center plate in each group alternately without touching either side (error) plate.

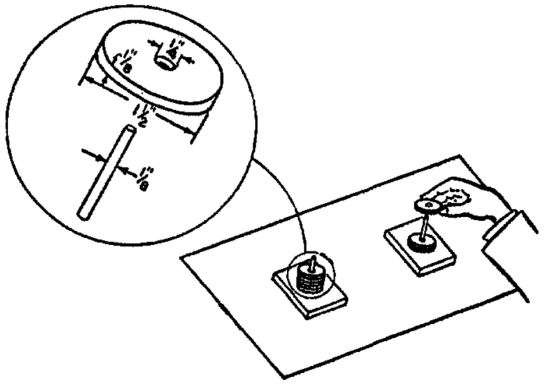


FIG. 2. Disc transfer apparatus. The task was to transfer eight washers one at a time from the right to the left pin. The inset gives the dimensions for the $W_s = \frac{1}{8}$ in. condition.

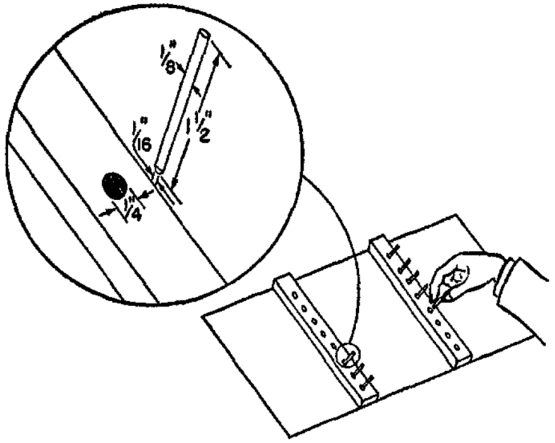
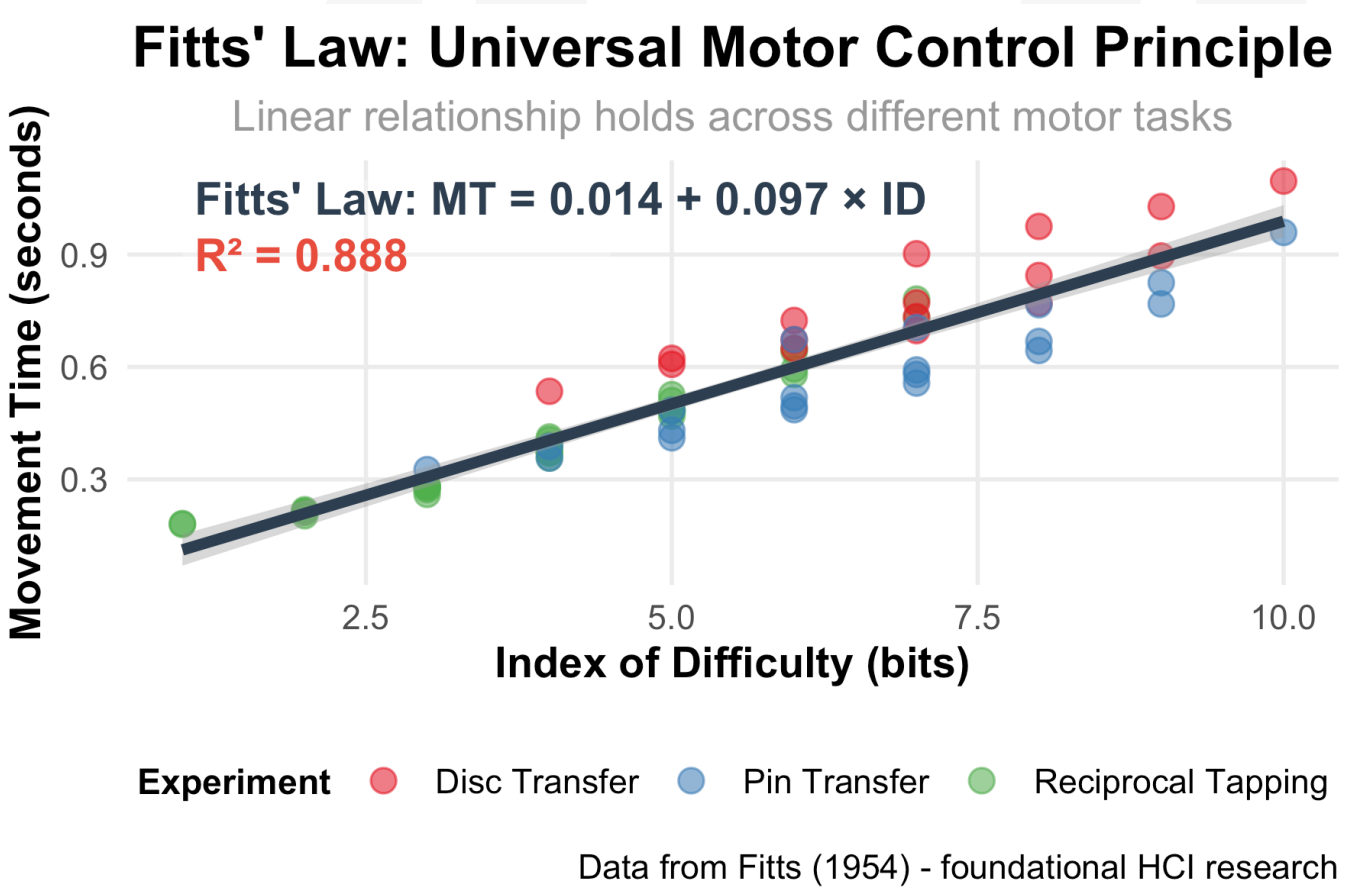


FIG. 3. Pin transfer apparatus. The task was to transfer eight pins one at a time from one set of holes to the other. The inset gives the dimensions of pins and holes for the $W_s = \frac{1}{8}$ in. condition.

Discovering Scientific Laws from Data



Data from Fitts (1954) - foundational HCI research

Discovering Scientific Laws from Data

Weber's Law
Passive dynamic walkers
Hick-Hyman Law
Stevens' Power Law
Signal Detection Theory
…

What do all of these have in common?

Weber's Law
Passive dynamic walkers
Hick-Hyman Law
Stevens' Power Law
Signal Detection Theory
…

What do all of these have in common?

Whether we are talking about cutting edge machine learning and artificial intelligence or Fechner tediously recoding his observations in Leipzig, the task is the same

We are using data to reason, mathematically, about behavior.

Take a moment to work with your neighbors

1. Write (or type) all the different kinds of modeling techniques you can think of

Take a moment to work with your neighbors

1. Write (or type) all the different kinds of modeling techniques you can think of
2. Think about each modeling technique and ask each other:
    1. Is this model interpretable?
    2. If no, why not?
    3. If yes, are there limits to the interpretation?

Models typically fall into two categories

Interpretable Models

- Explicit Mathematical Models
  - Haken-Kelso-Bunz Model
  - Sine-circle map
  - Passive Dynamic Walker (compass gait)
- Multiple regression
- Structural equation models

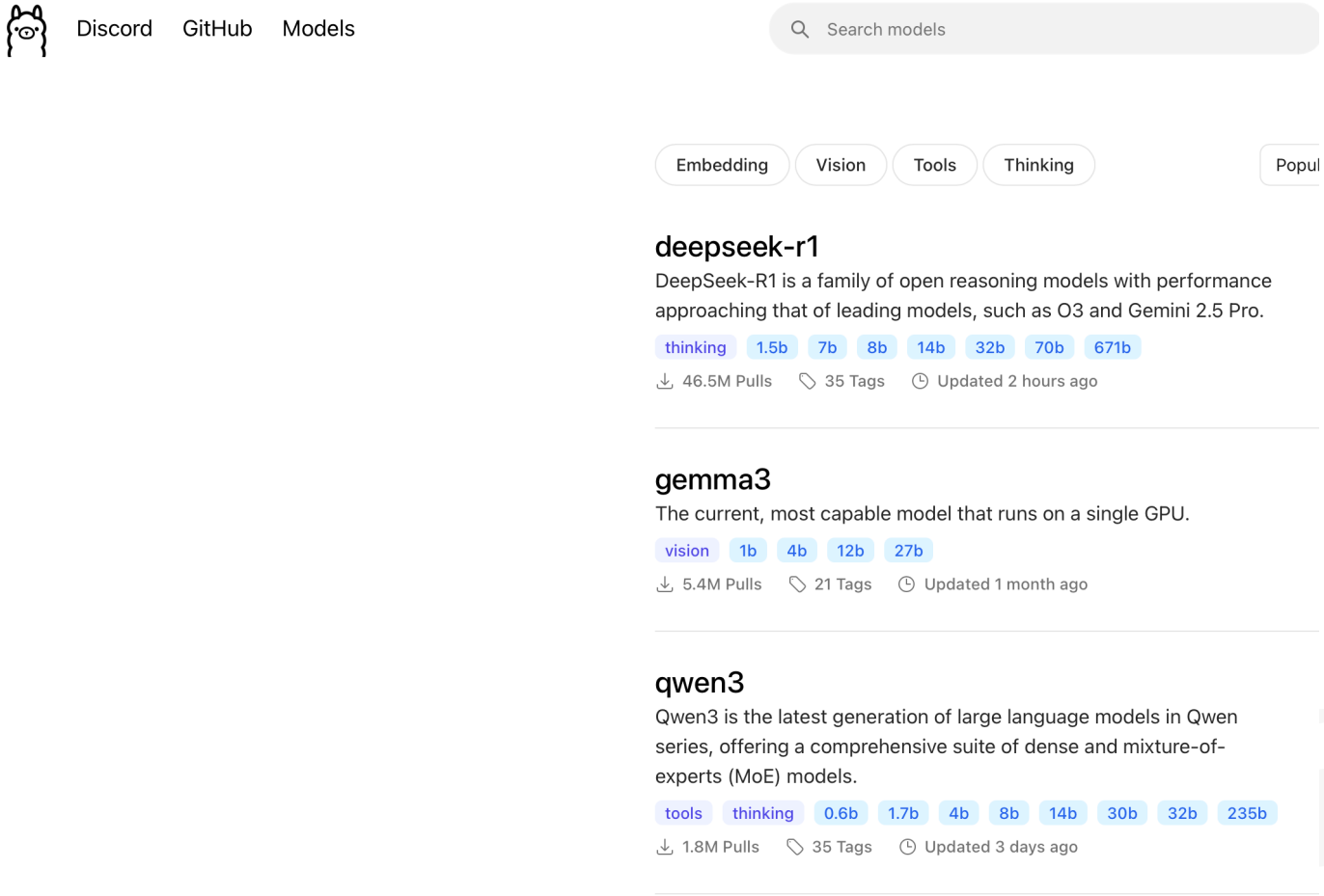
Non-Interpretable models

- Support Vector Machines*
- Random Forests
- Naïve Bayes
- Deep Neural Nets
  - MLPs
  - CNNs
  - RNNs
  - Modern LLMs

Just how uninterpretable are deep neural nets?

Model	2D-CNN	3D-CNN	Semi-CNN	
	Params	Params	Pre-Trained Params	Total Params
VGG-16	134.7 M	179.1 M	5.3 M	82.2 M
ResNet-18	11.4 M	33.3 M	0.4 M	31.7 M
ResNet-34	21.5 M	63.6 M	0.8 M	60.5 M
ResNet-50	23.9 M	46.4 M	0.9 M	45.8 M
ResNet-101	42.8 M	85.5 M	0.9 M	84.8 M
ResNet-152	58.5 M	117.6 M	1.4 M	115.6 M
DenseNet-121	7.2 M	11.4 M	0.8 M	10.4 M
DenseNet-169	12.8 M	18.8 M	0.8 M	17.9 M

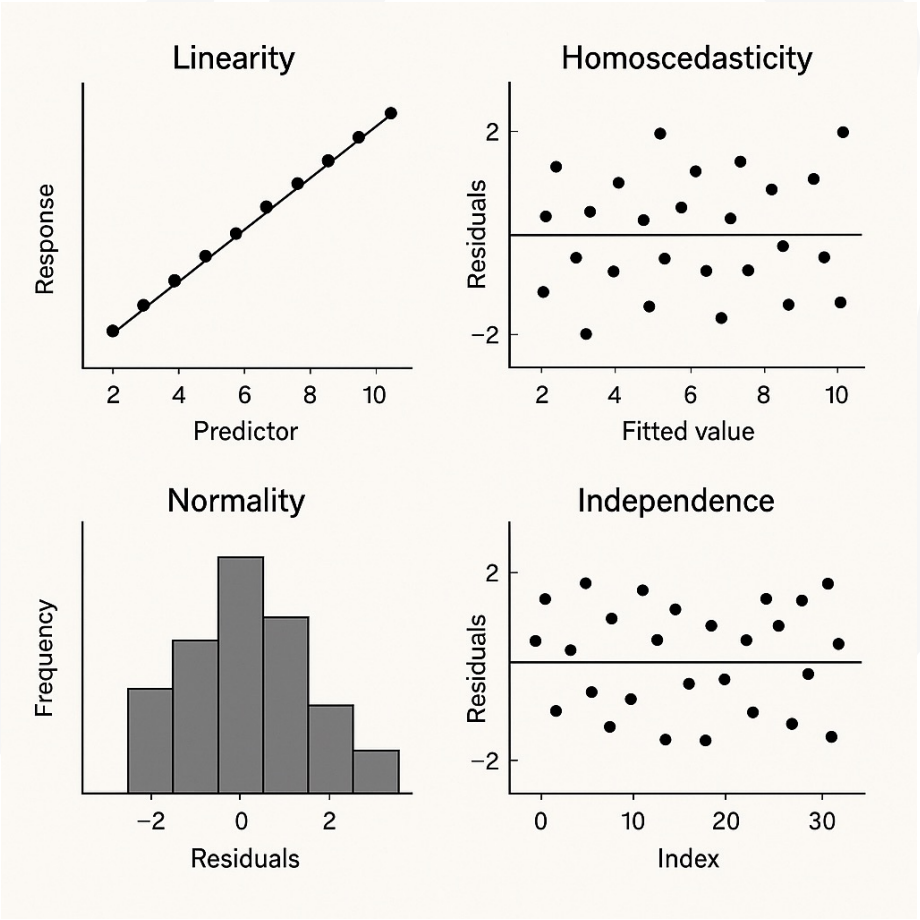
Just how uninterpretable are deep neural nets?



Regression models are generally interpretable, but…

Who remembers their regression assumptions?

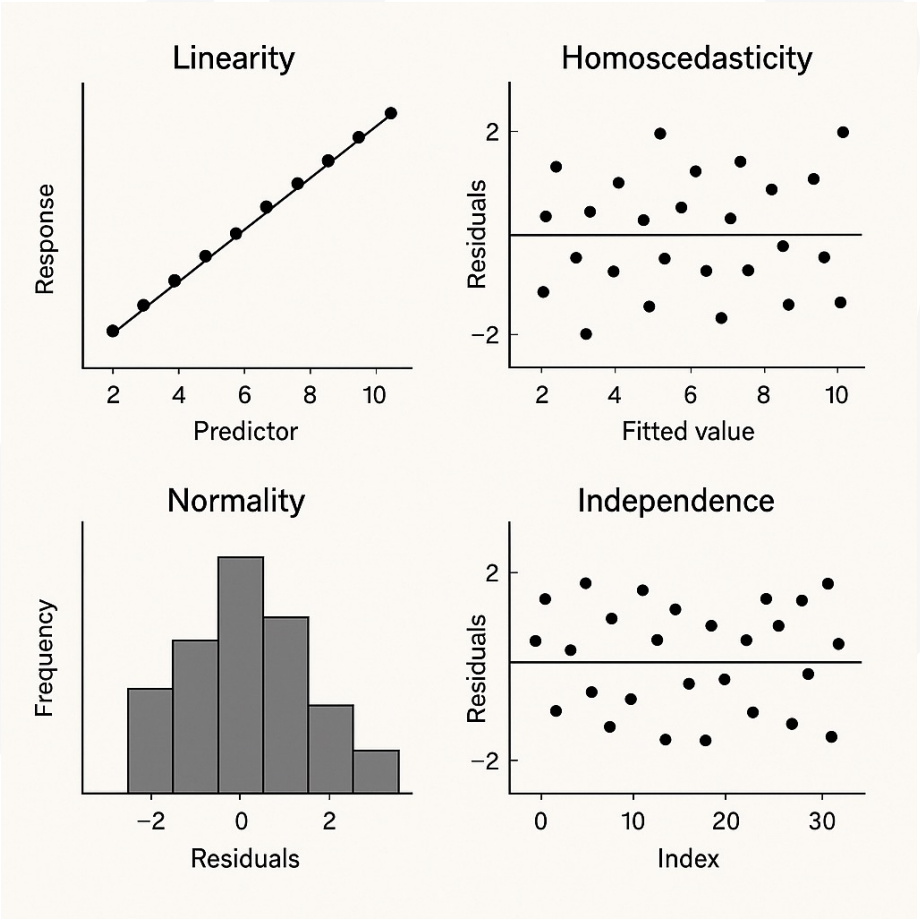
- Linearity
- Independence
- Normality
- Homoscedasticity



Regression models are generally interpretable, but…

Who remembers their regression assumptions?

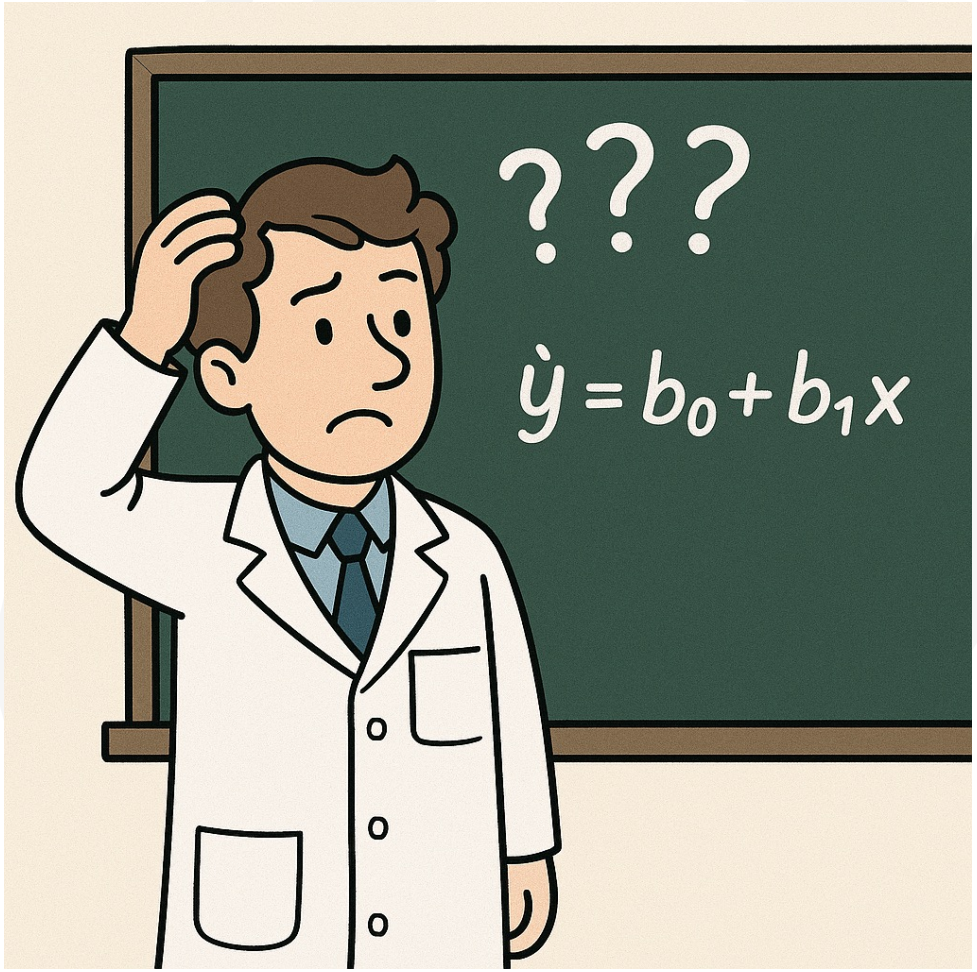
- Linearity
- Independence
- Normality
- Homoscedasticity
- **Correct Specification**



Regression models are generally interpretable, but…

They require us to specify our models in advance along with any transformations of our data.

What if we don't know the functional form our models should take?



Symbolic Regression as a Solution

Symbolic regression refers to to a group of techniques that discover equations from data

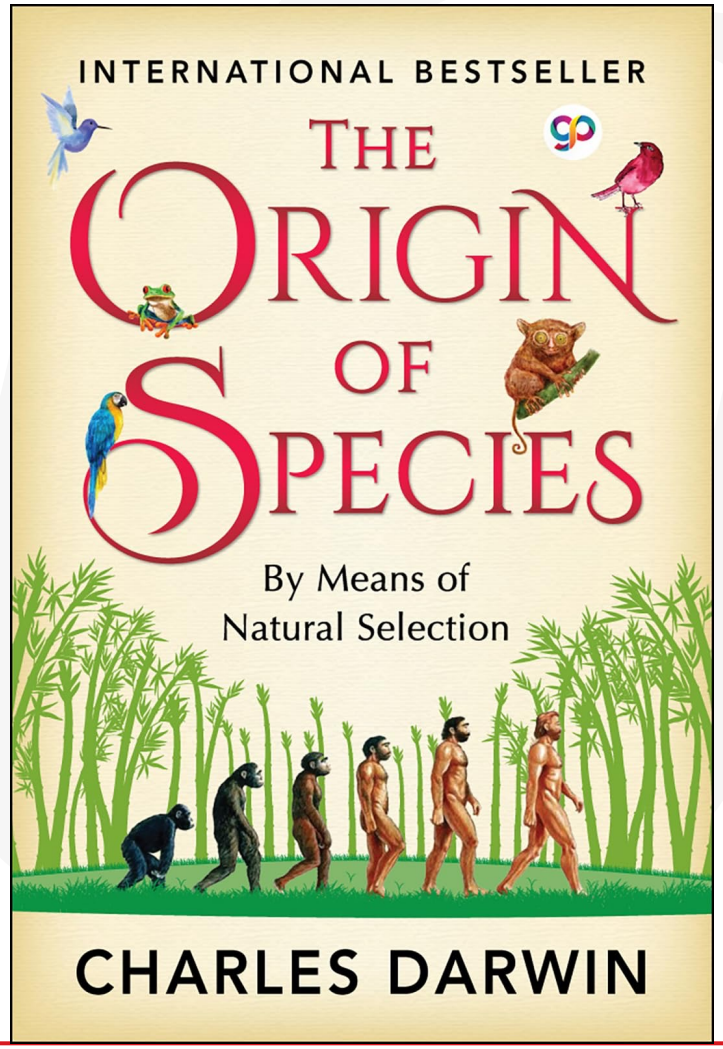
Symbolic regression can (although it needn't) incorporate domain knowledge (e.g., maintain units, physical constraints)

There are many variants. We will focus on one that involves genetic programming.

Genetic Programming

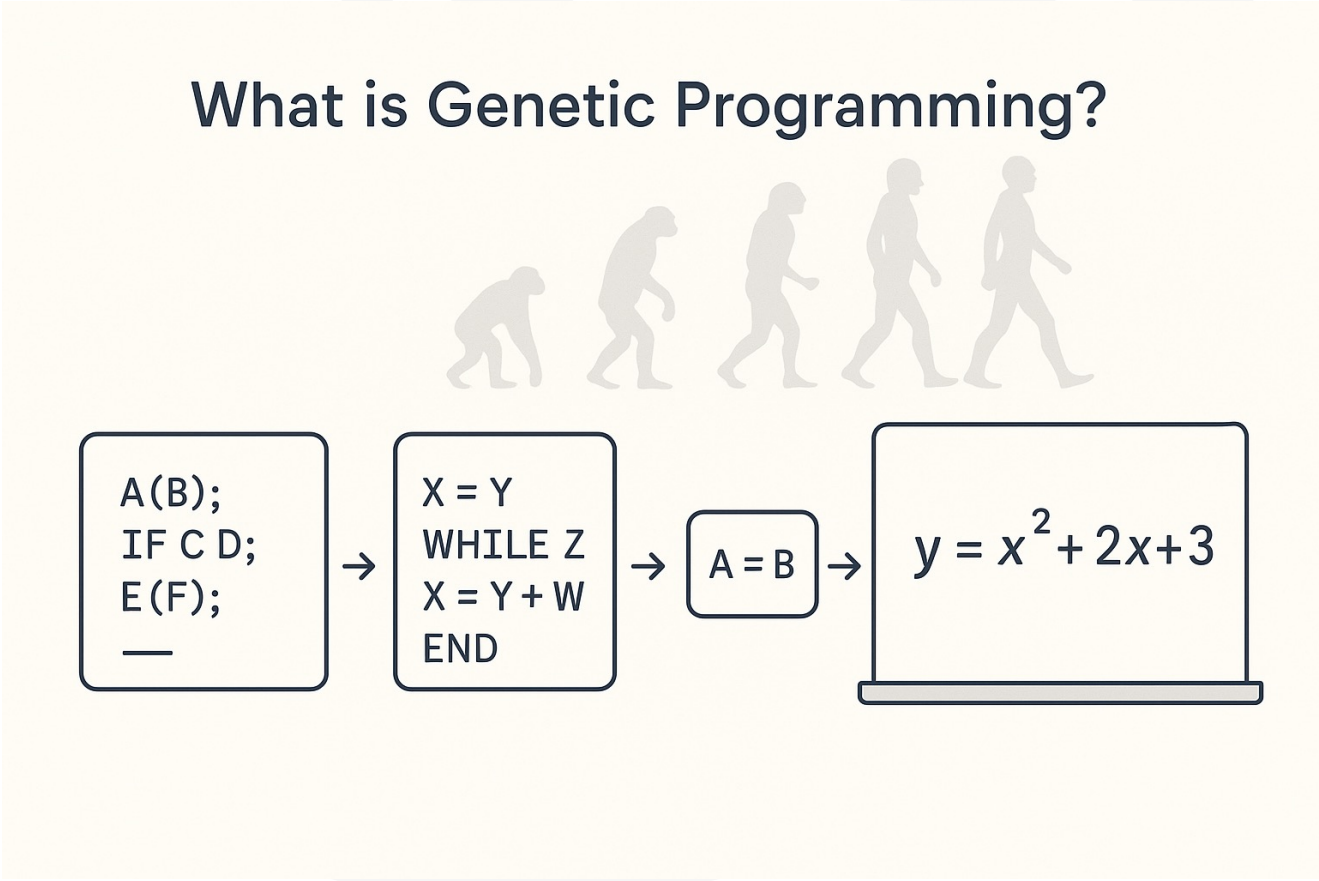
What is genetic programming?

- Evolution-inspired method for automatically writing computer programs.
- Programs evolve over generations using selection and variation.
- A type of evolutionary algorithm, like genetic algorithms but for code.



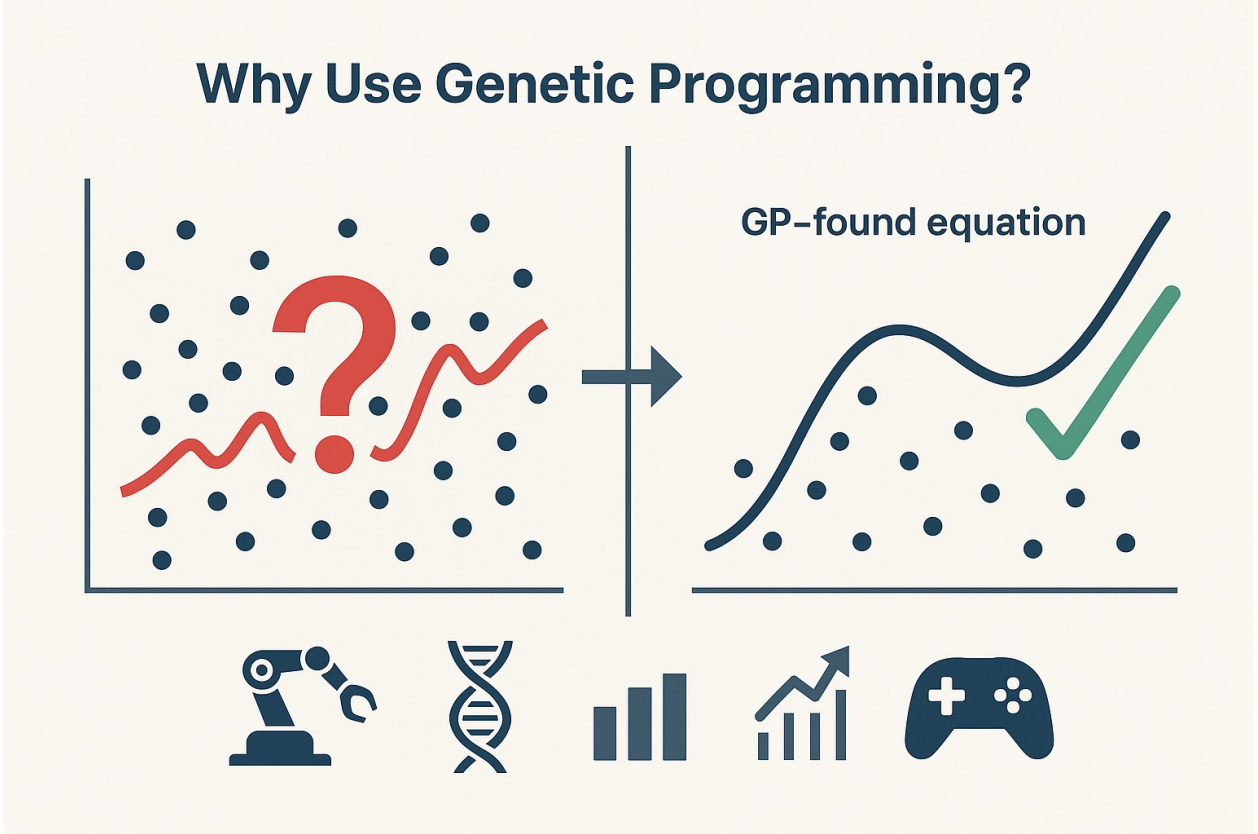
What is genetic programming?

- Evolution-inspired method for automatically writing computer programs.
- Programs evolve over generations using selection and variation.
- A type of evolutionary algorithm, like genetic algorithms but for code.



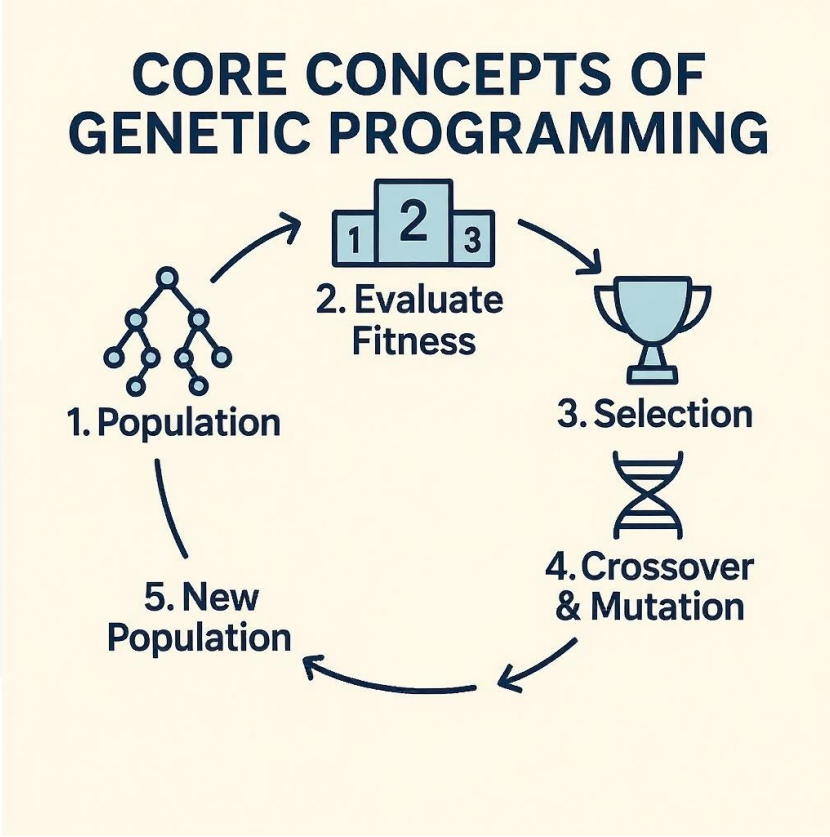
Why should you use genetic programming?

- Discovers solutions when we can't write rules ourselves.
- Handles complex, nonlinear relationships.
- Produces human-readable programs or equations.
- Useful when the search space is huge or poorly understood.



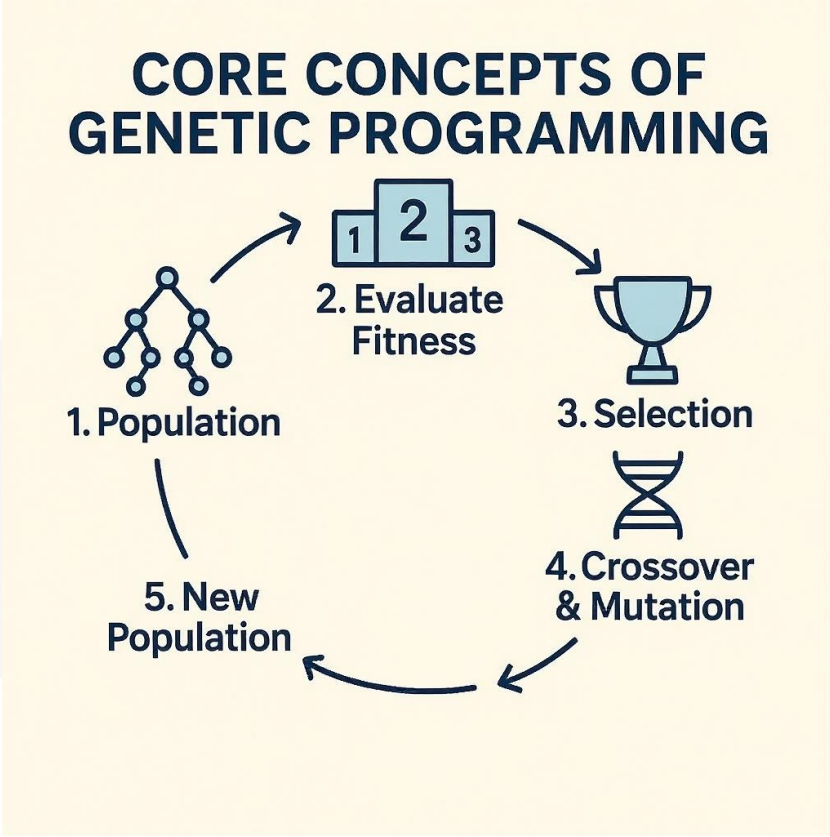
Core Concepts of Genetic Programming

- Individual = candidate program (tree of operations).



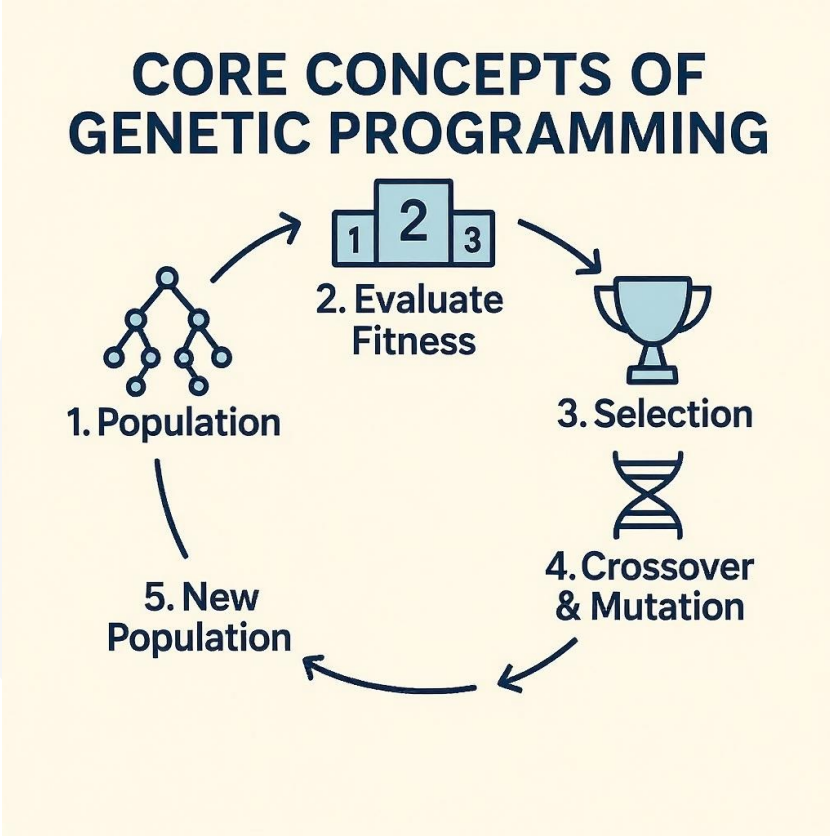
Core Concepts of Genetic Programming

- Individual = candidate program (tree of operations).
- Population = many individuals evaluated each generation.



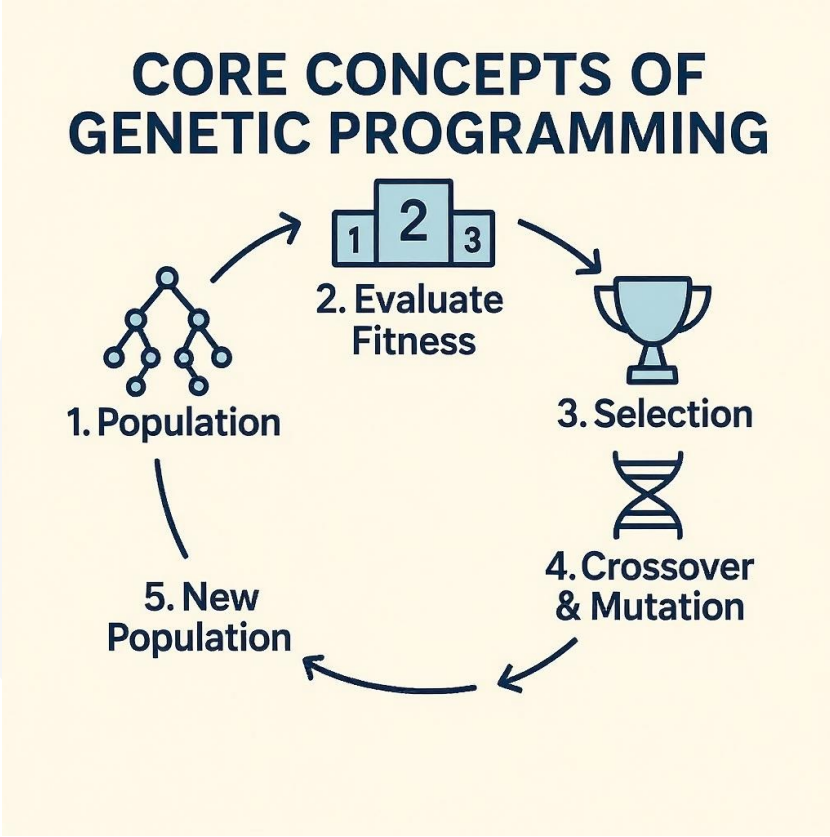
Core Concepts of Genetic Programming

- Individual = candidate program (tree of operations).
- Population = many individuals evaluated each generation.
- Fitness function = how well a program solves the task.



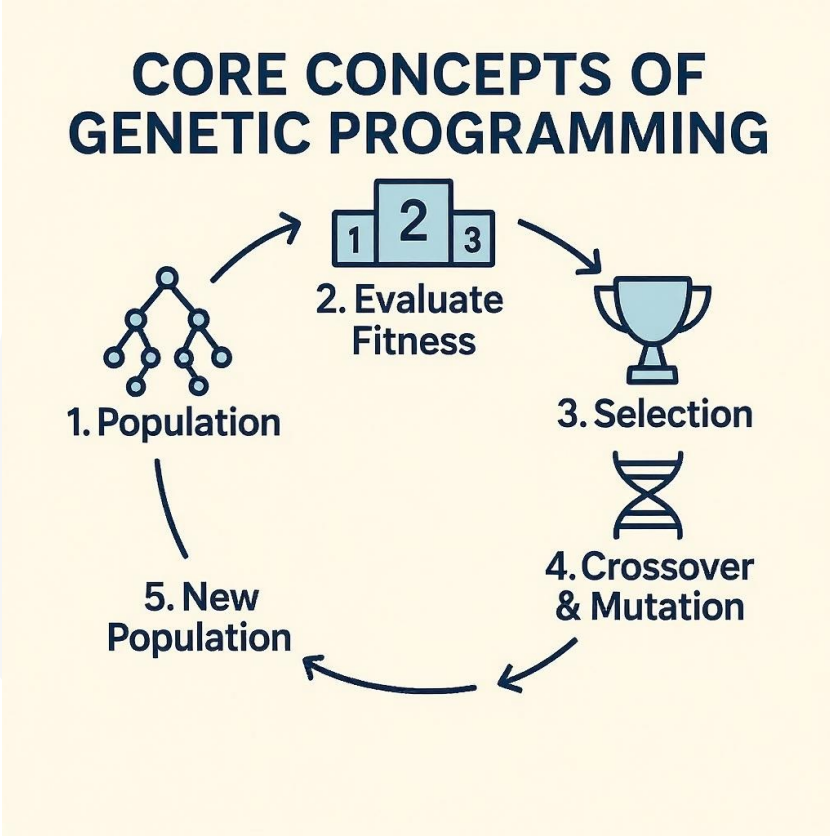
Core Concepts of Genetic Programming

- Individual = candidate program (tree of operations).
- Population = many individuals evaluated each generation.
- Fitness function = how well a program solves the task.
- Selection = choose better individuals to be parents.



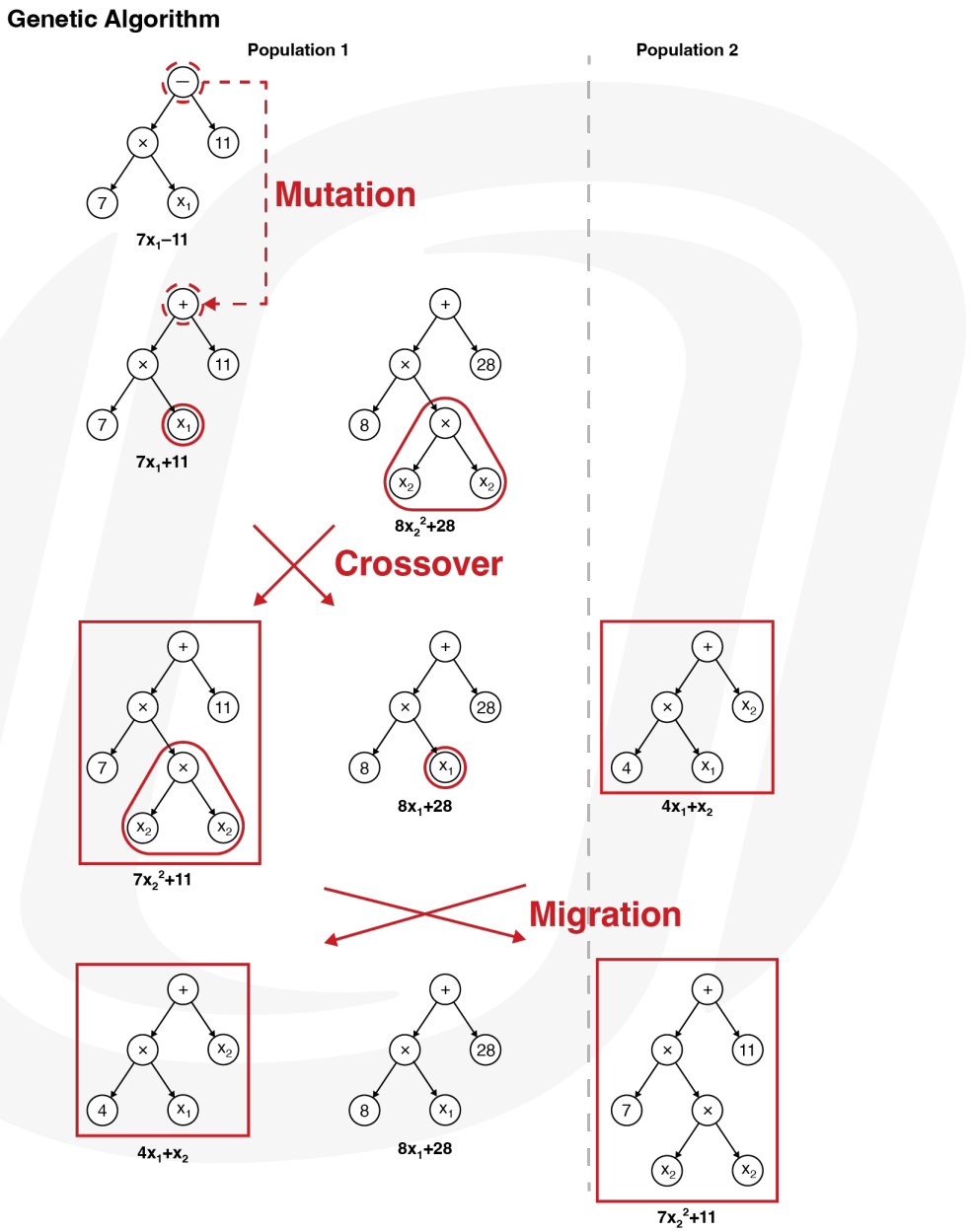
Core Concepts of Genetic Programming

- Individual = candidate program (tree of operations).
- Population = many individuals evaluated each generation.
- Fitness function = how well a program solves the task.
- Selection = choose better individuals to be parents.
- Crossover & mutation = create new programs by recombining or altering parents.



Representation and operators

- Typical representation: expression trees (functions are internal nodes, inputs/constants are leaves).
- Crossover: swap sub-trees between parents.
- Mutation: randomly change a node, constant, or sub-tree.
- Keeps syntactically correct programs through type constraints.



Let’s take a break and then…
Game Time!

Card Game

1. Piles of cards will be prepared.
    - Each pile of card will contain each gene from the preset gene pool
2. Players will randomly select cards and create equations that are sizes equal or less than the max equation size.
    - Equations should follow algebraic rules.
3. Each player's equation will be evaluated.
    - Best equations at each equation size will be stored.

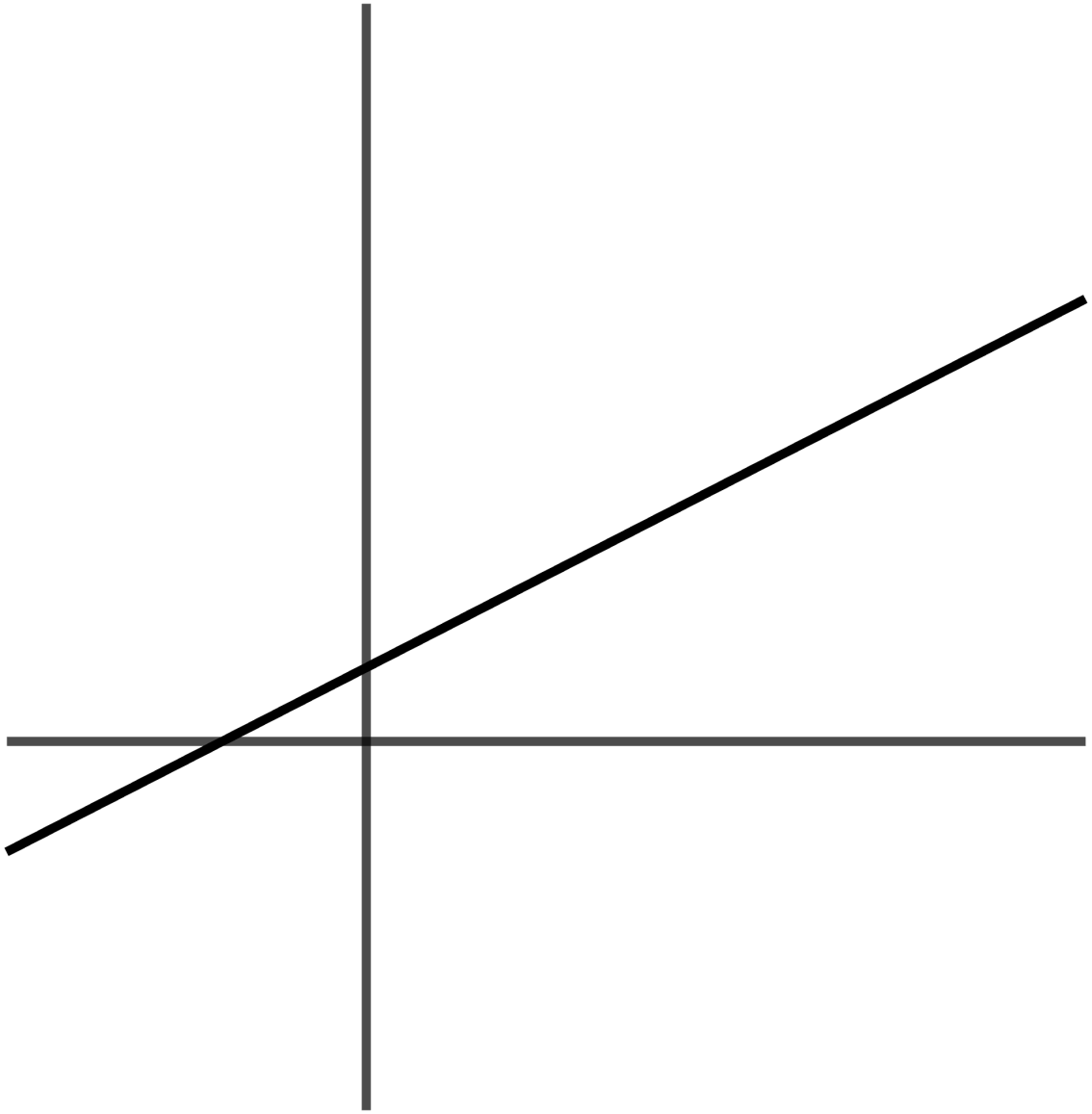
Card Game

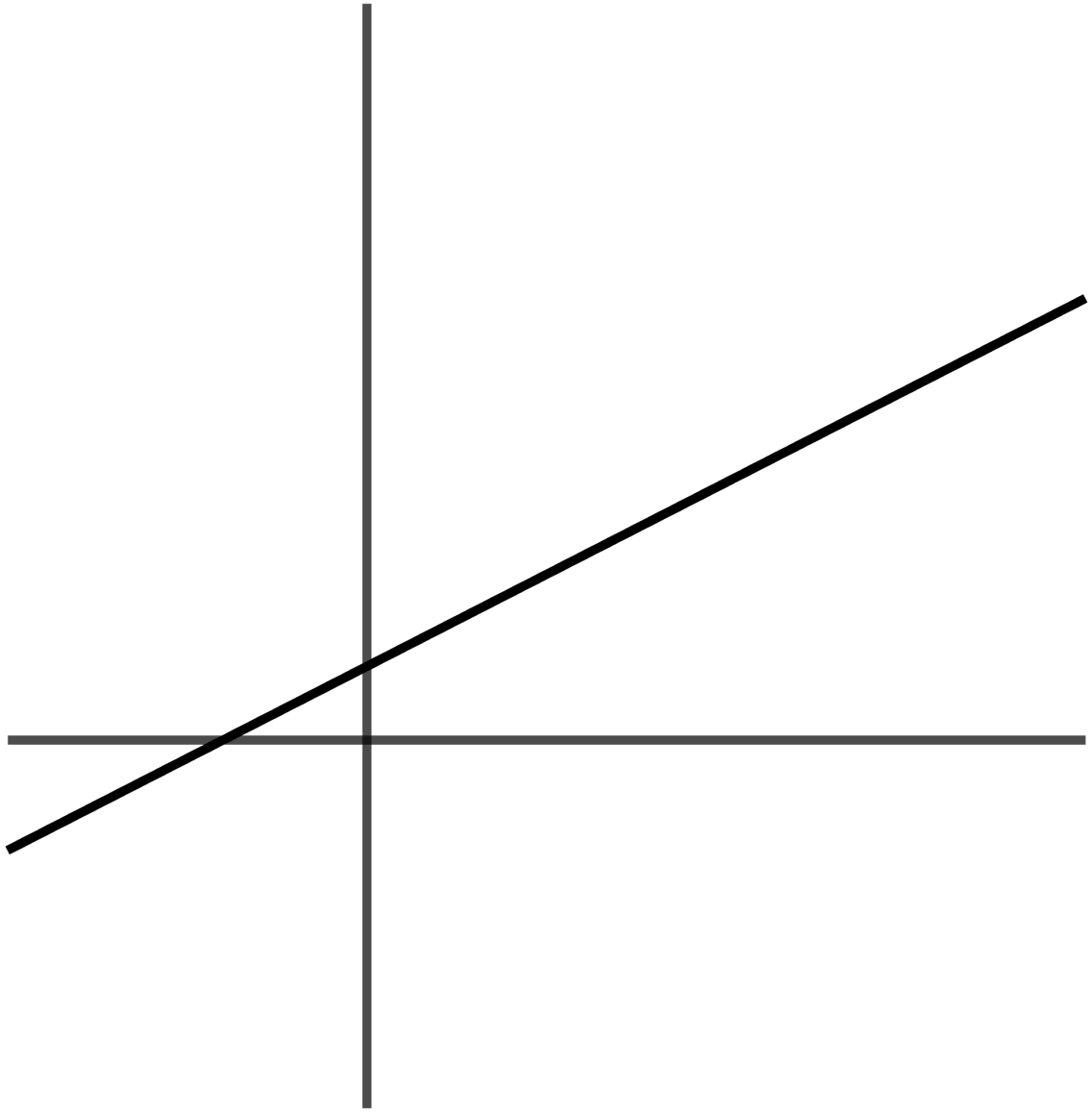
4. The best two will survive and the worst two will die off.
5. The best two will be reproduced with probabilistic crossover, then mutation.
    - Crossover and mutation will happen based on coin flips.
6. Steps 2-5 will be repeated 3 times.

Card Game

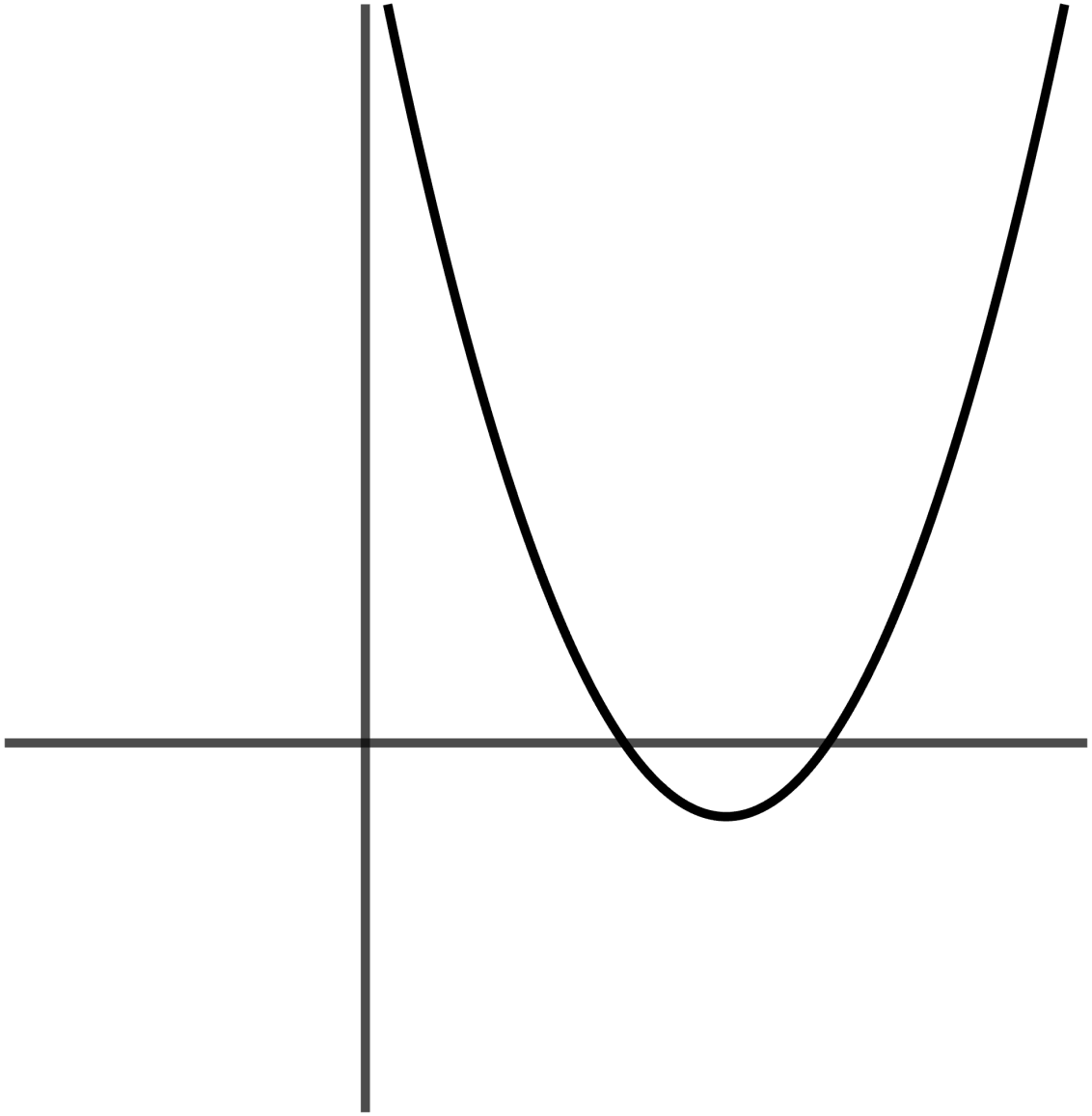
7. One of the players on the table will be chosen randomly and move over to the next table.
8. Step 6-7 will be repeated 3 times.
9. Final equations will be evaluated.
10. Best Equations at each equation size becomes the Pareto front.

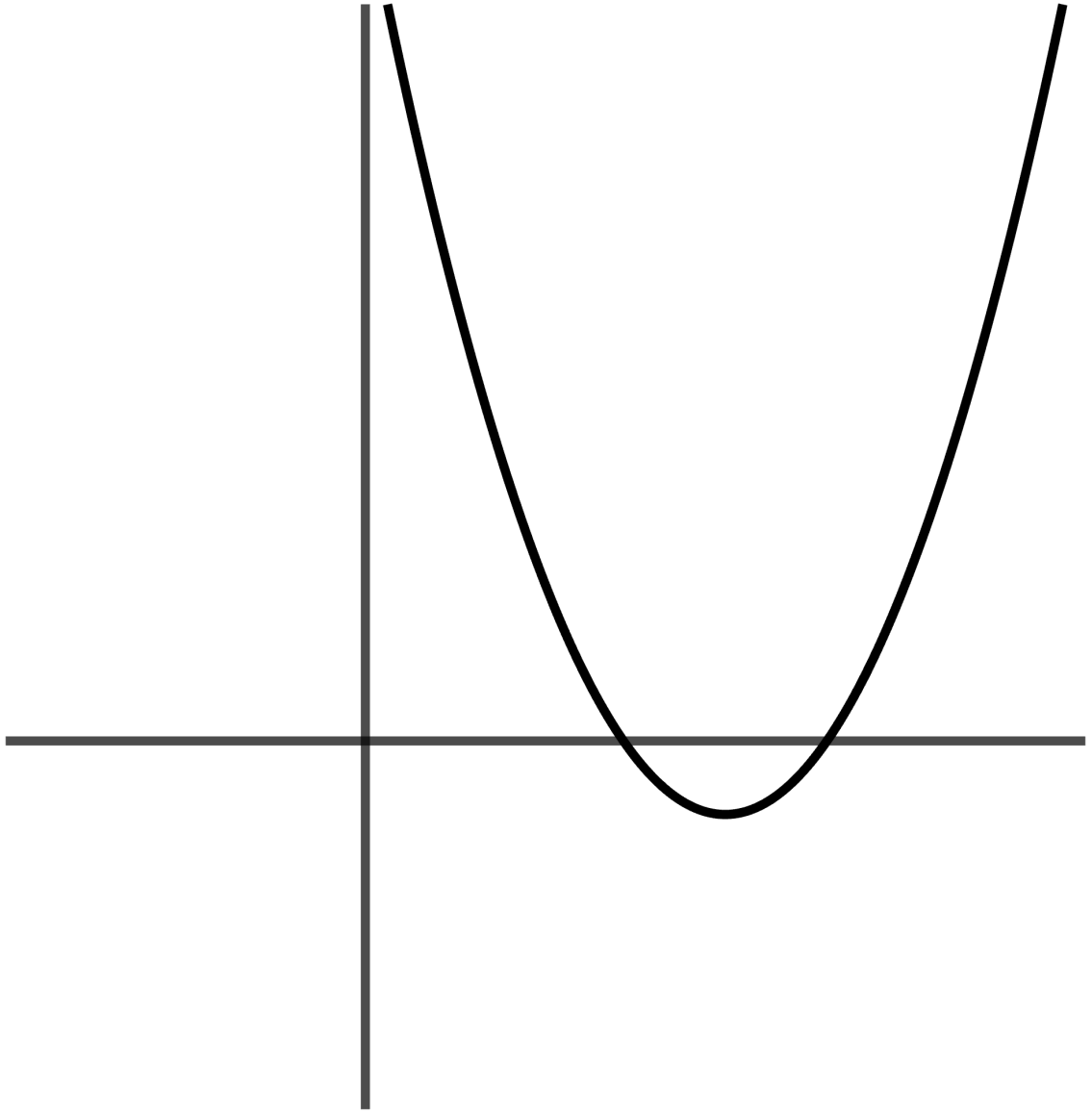
Symbolic Regression



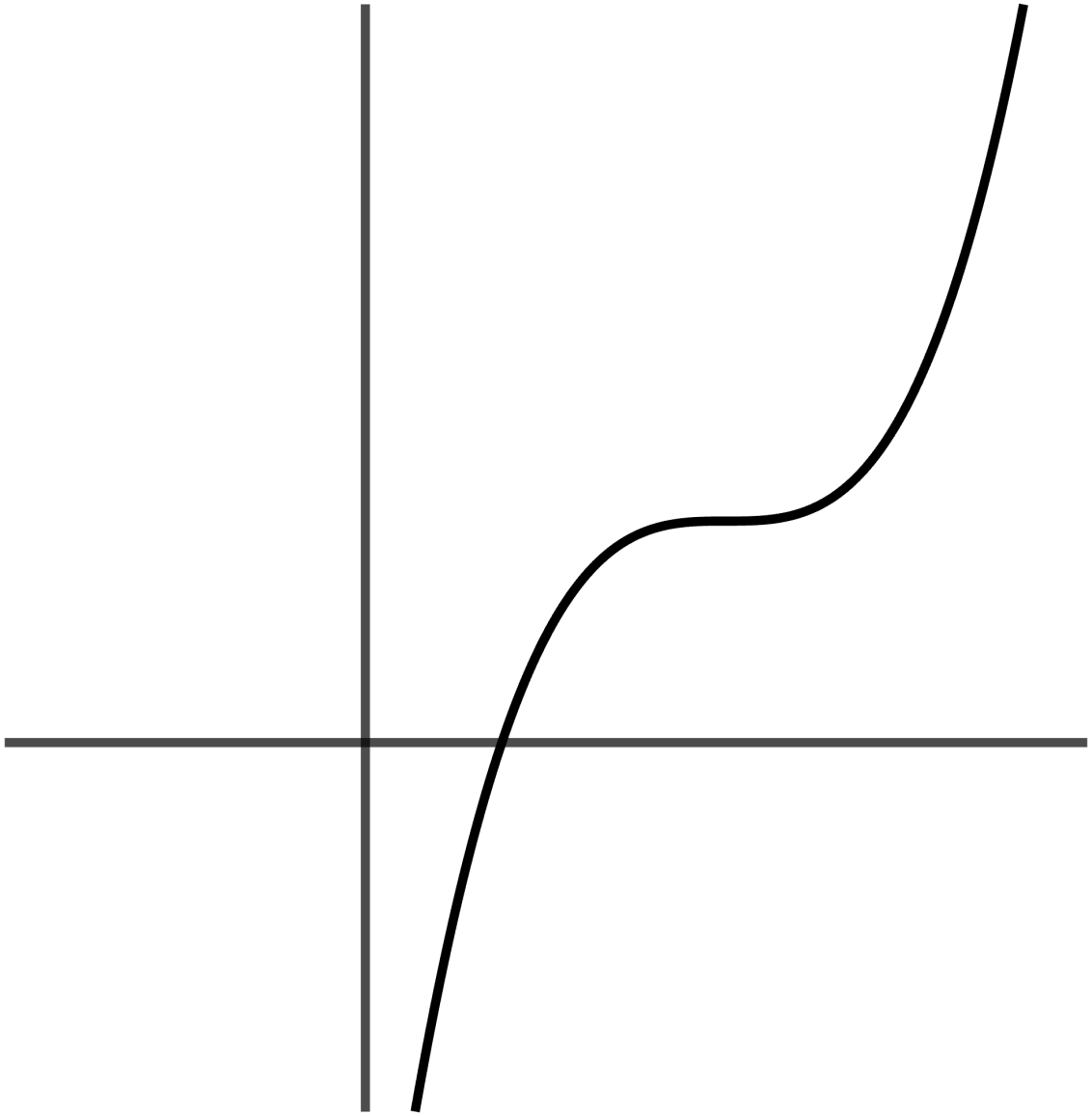


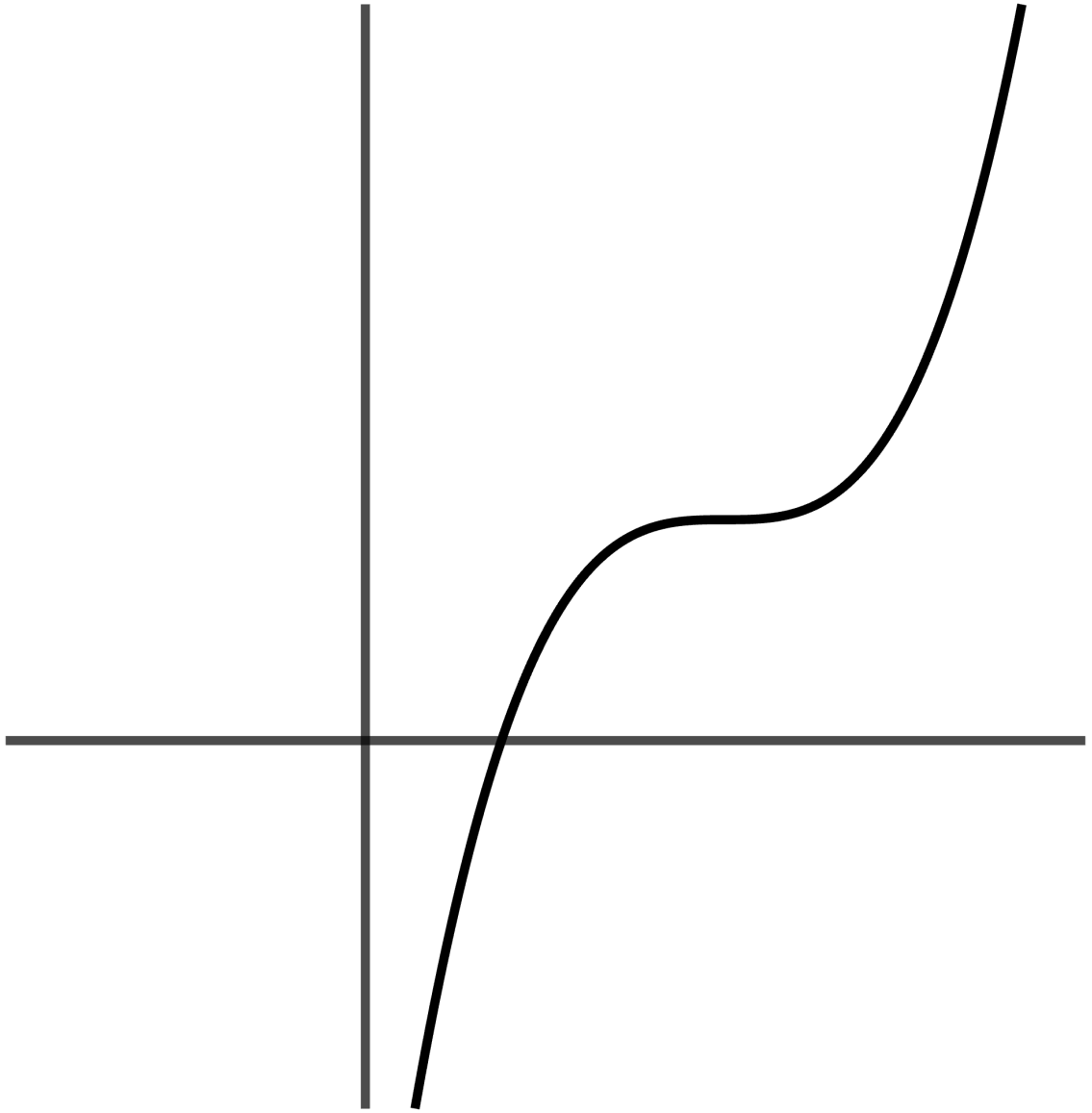
$$y = ax + b$$



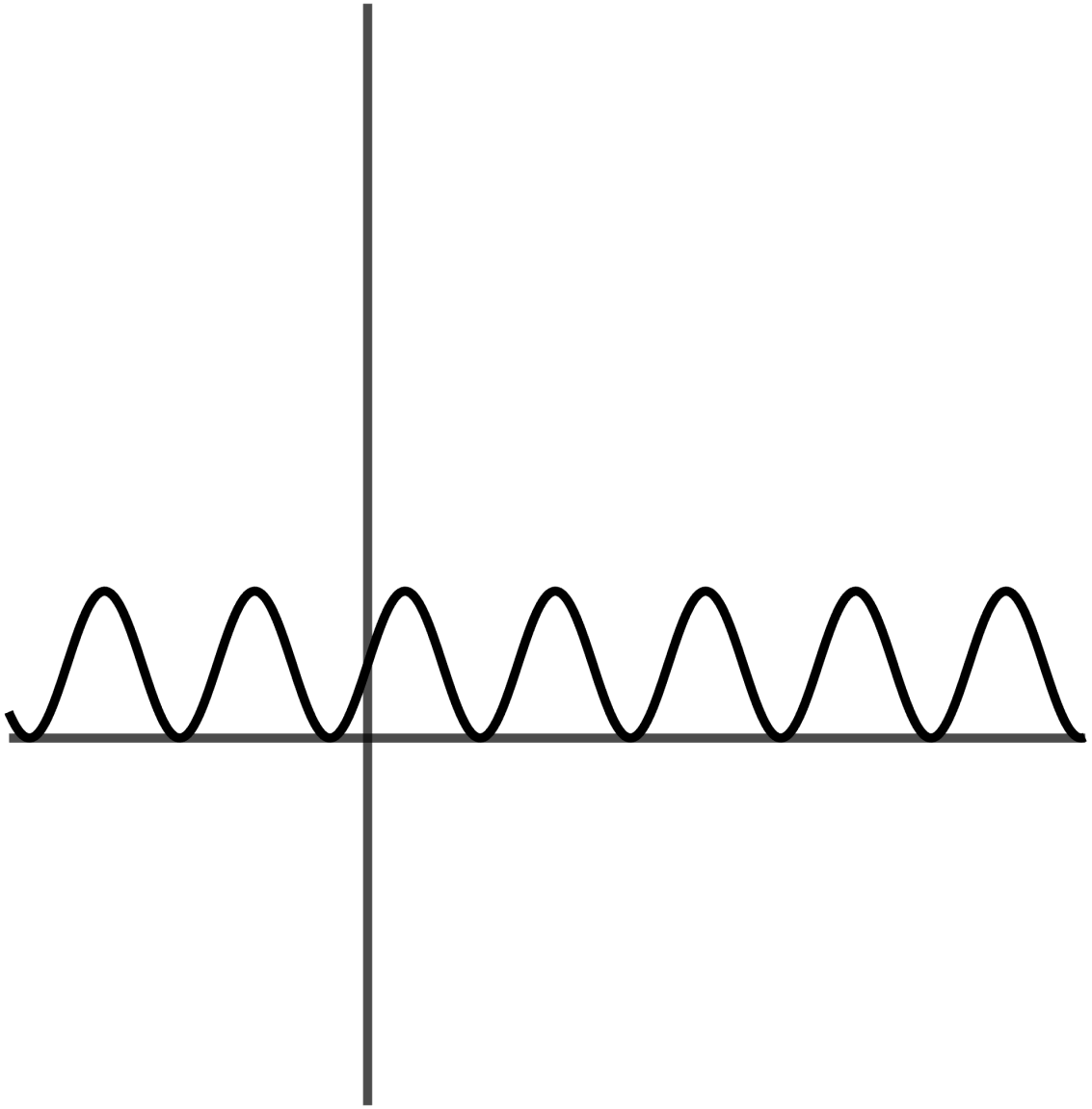


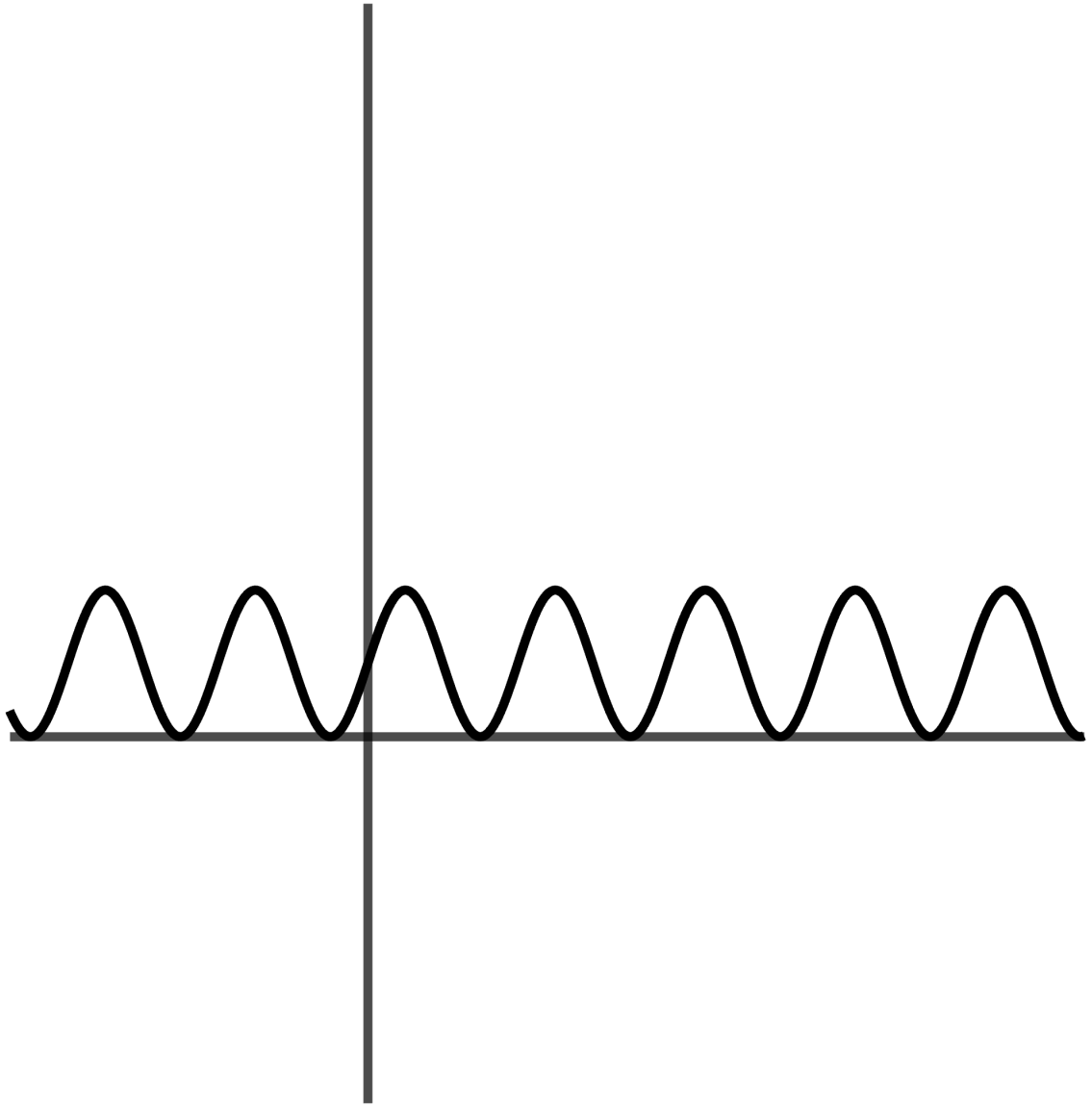
$$y = ax^2 + bx + c$$



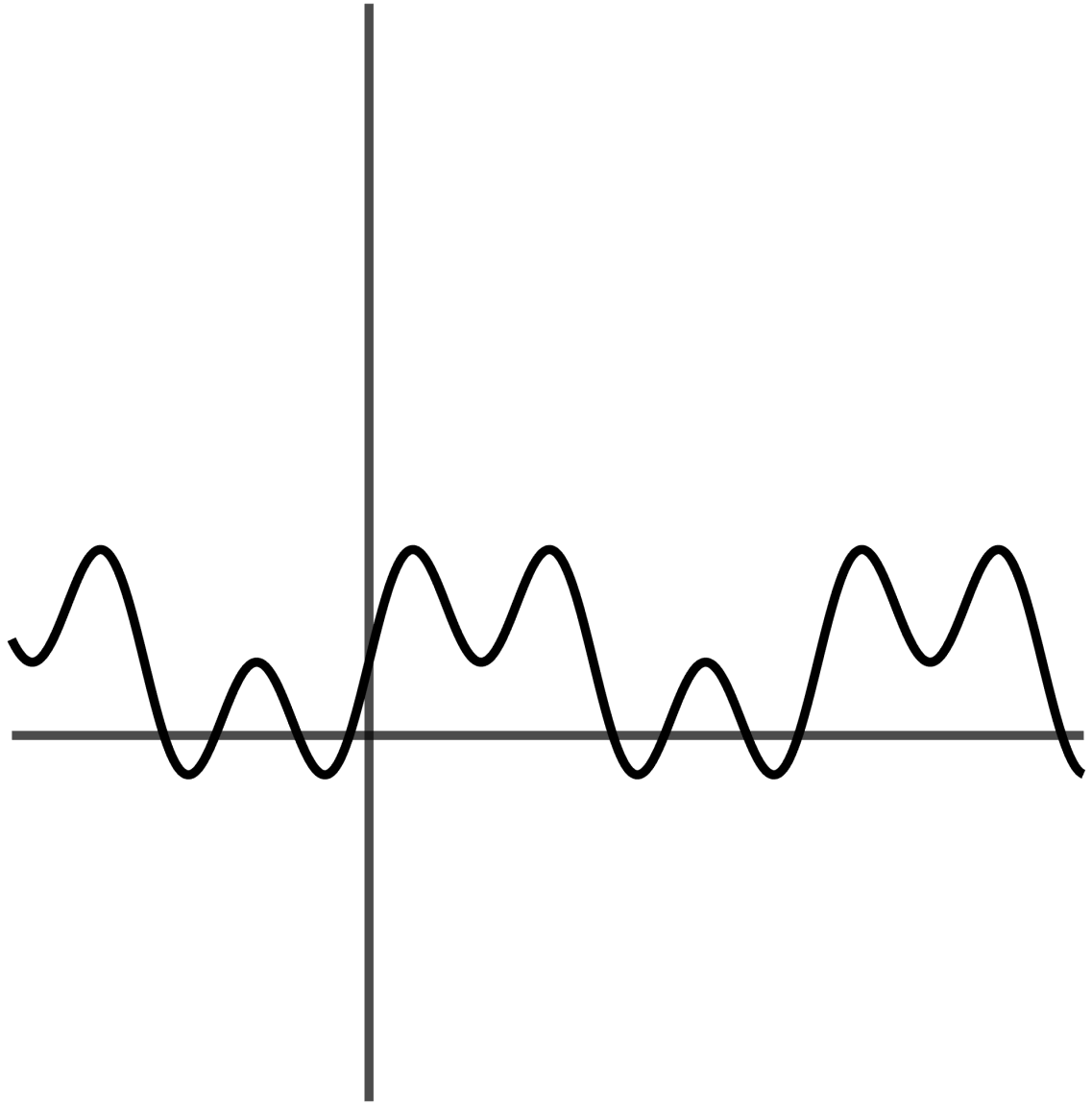


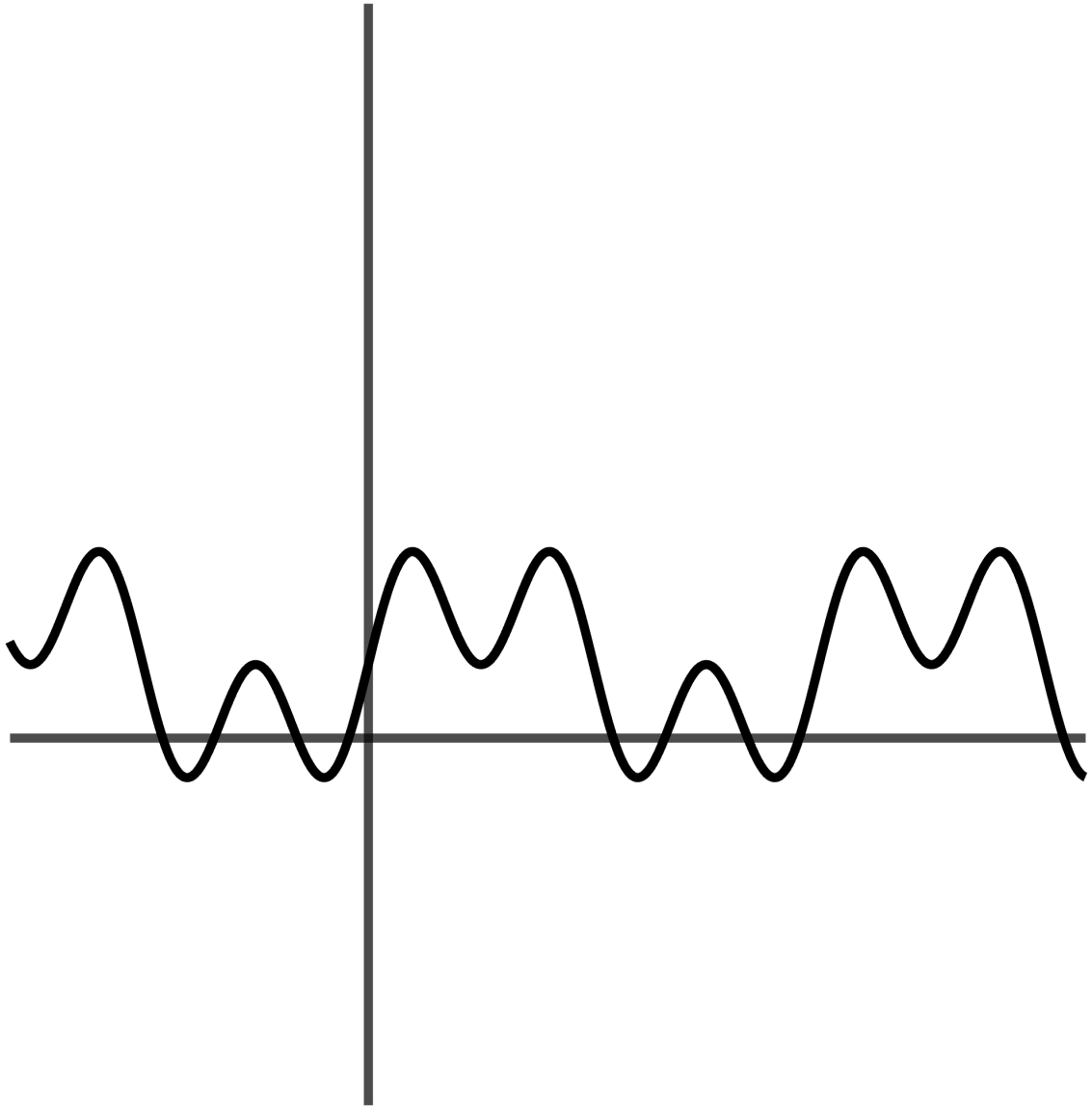
$$y = ax^3 + bx^2 + cx + d$$



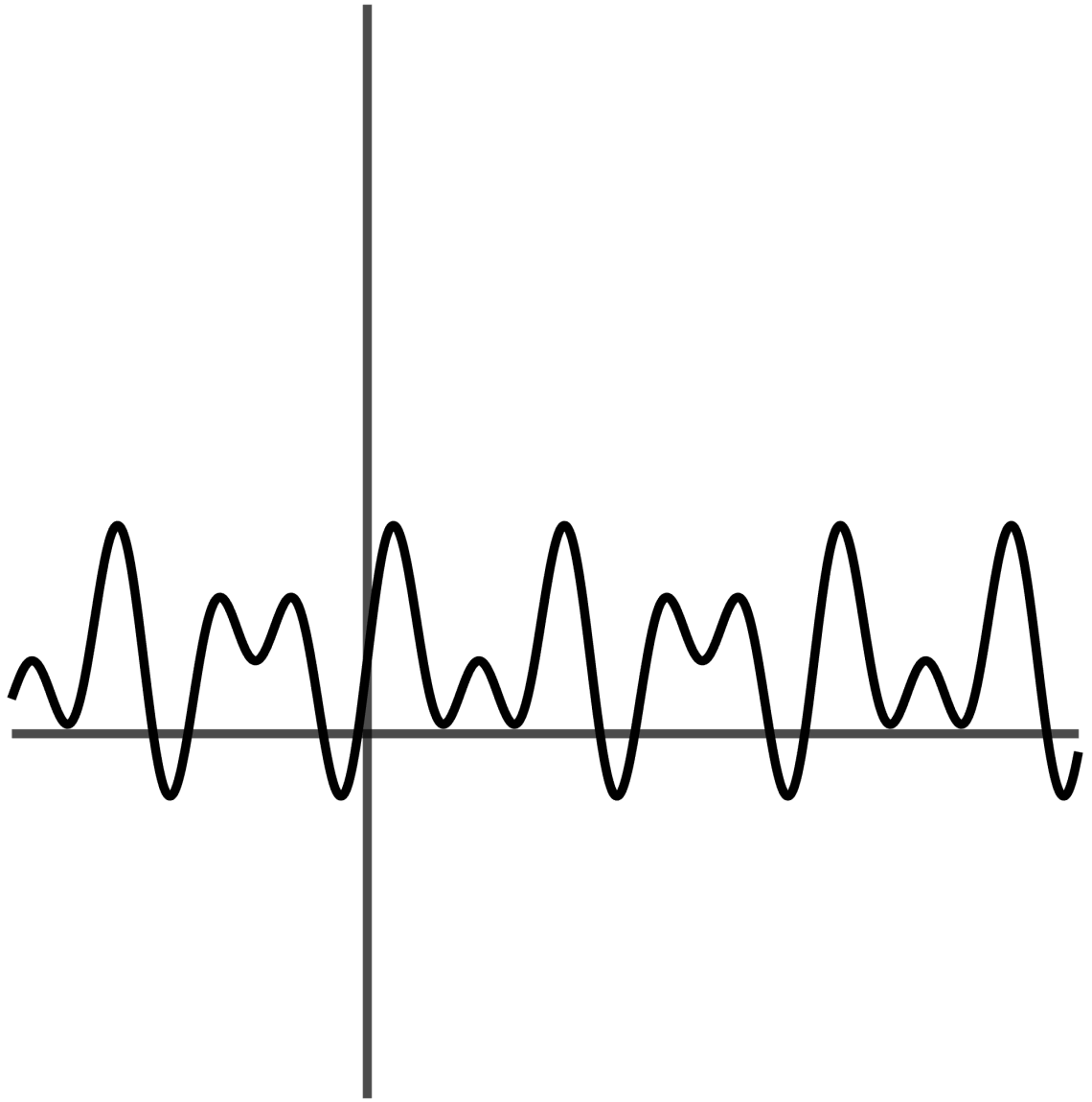


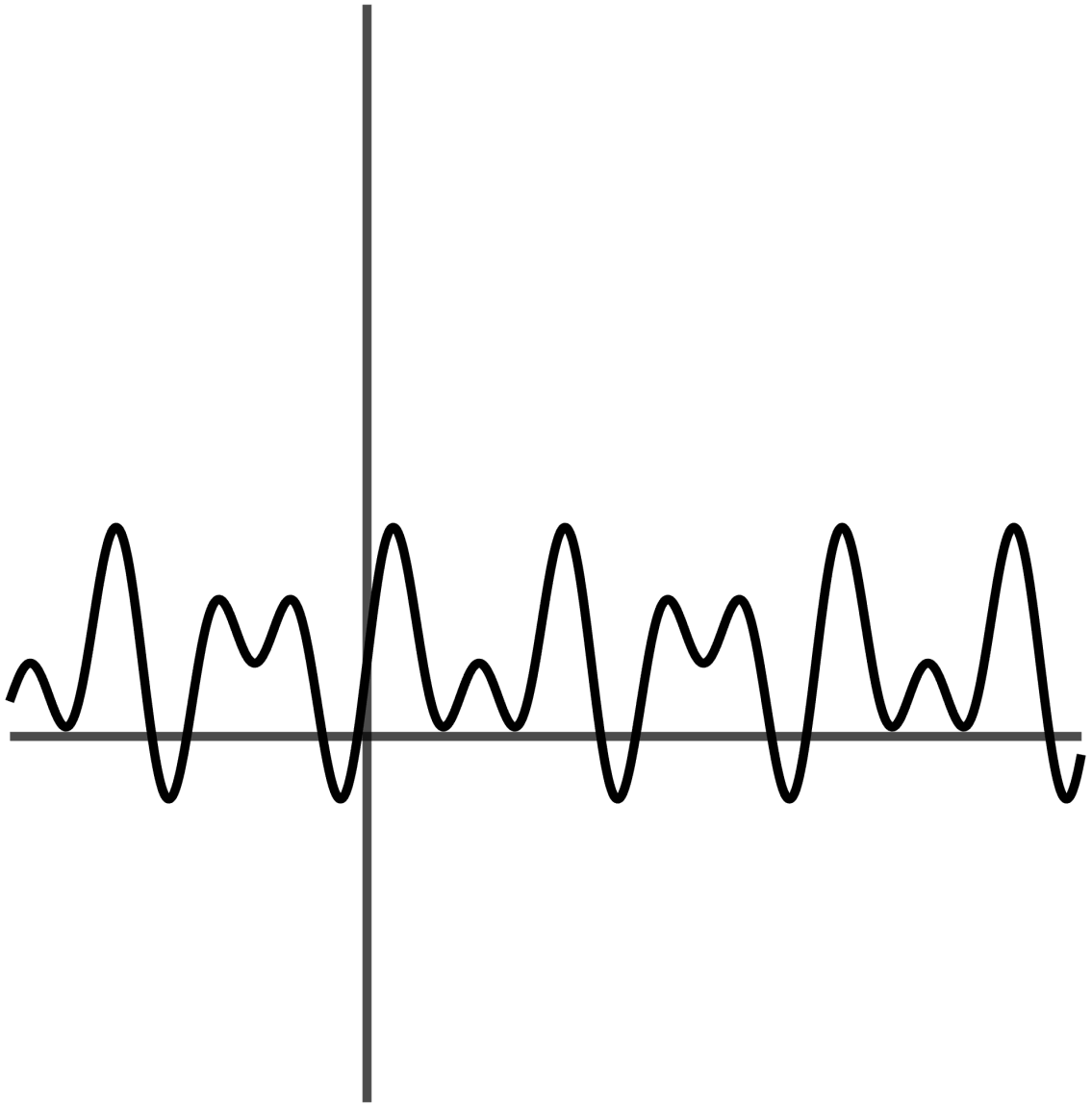
$$y = \sin(ax) + 1$$



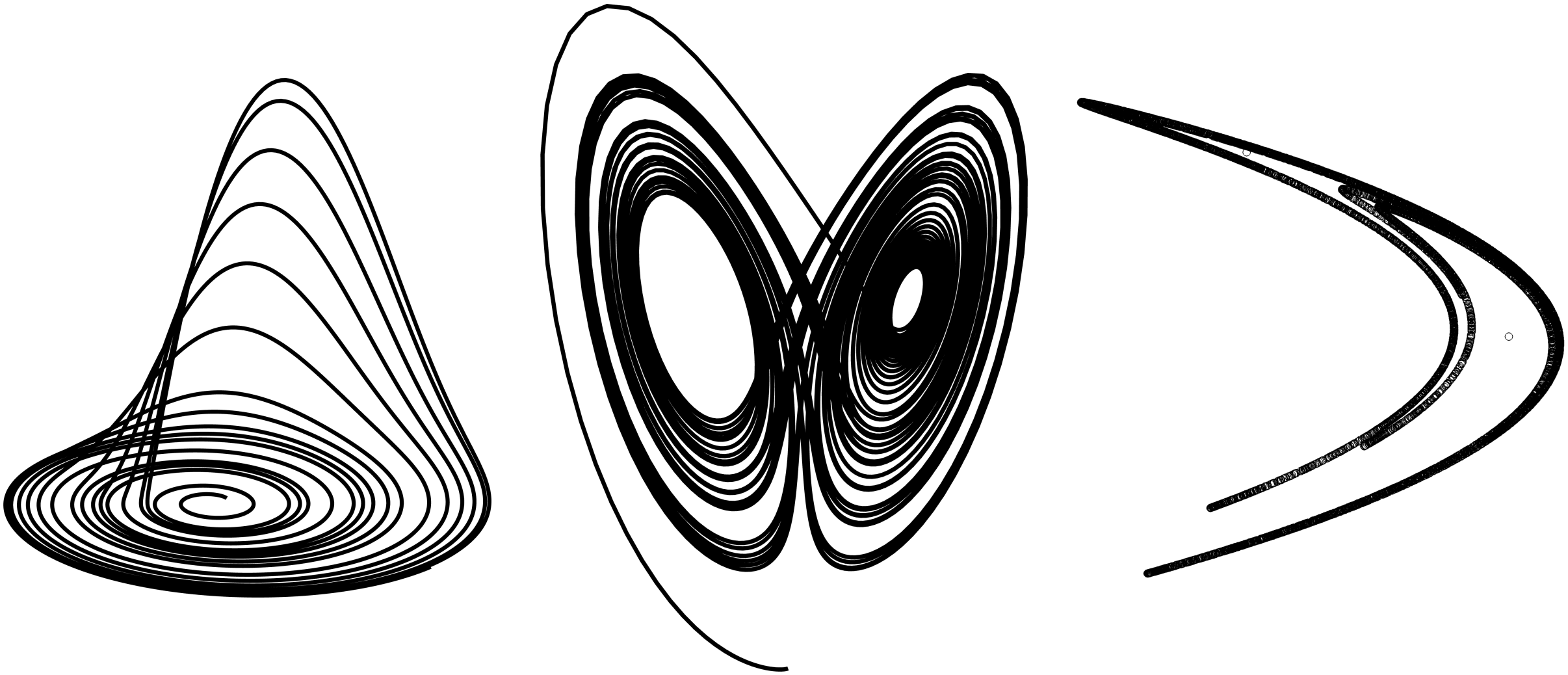


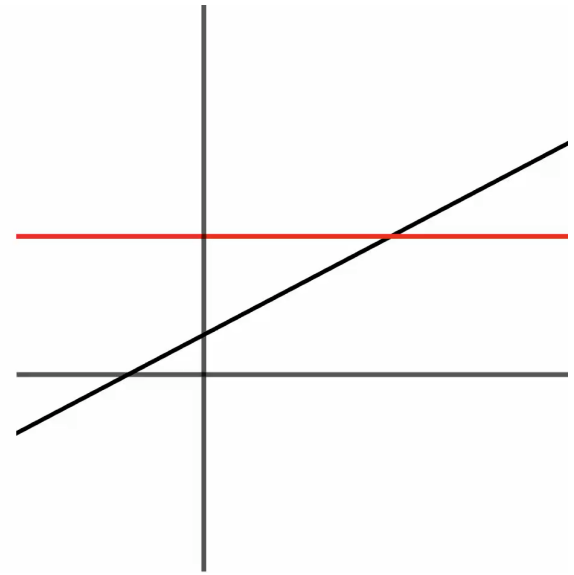
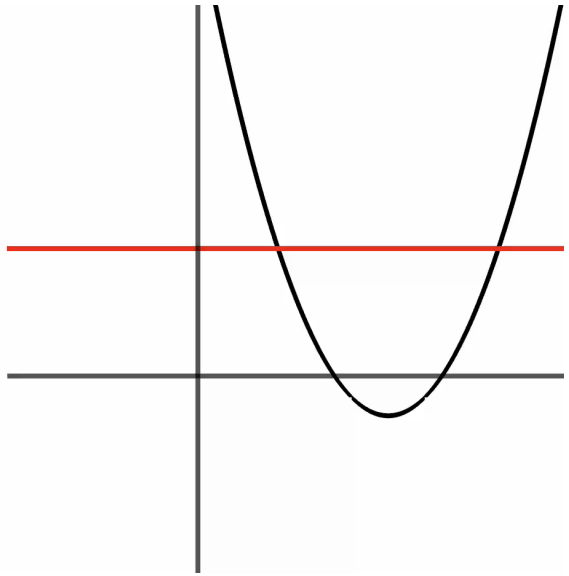
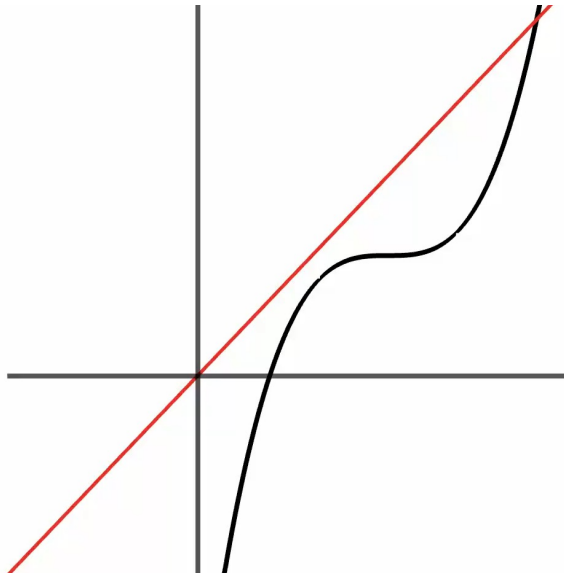
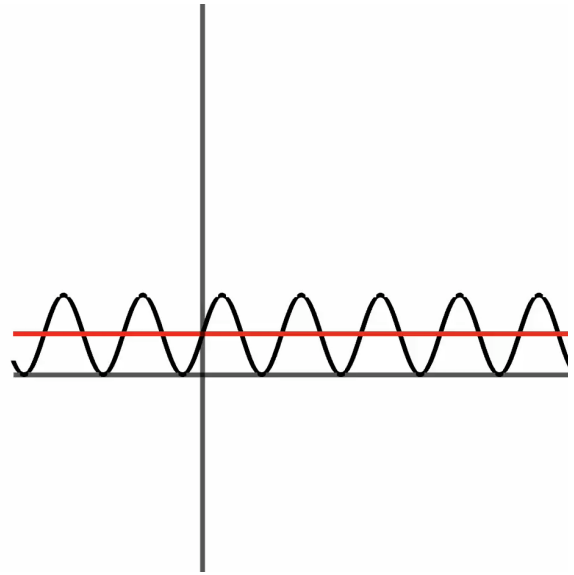
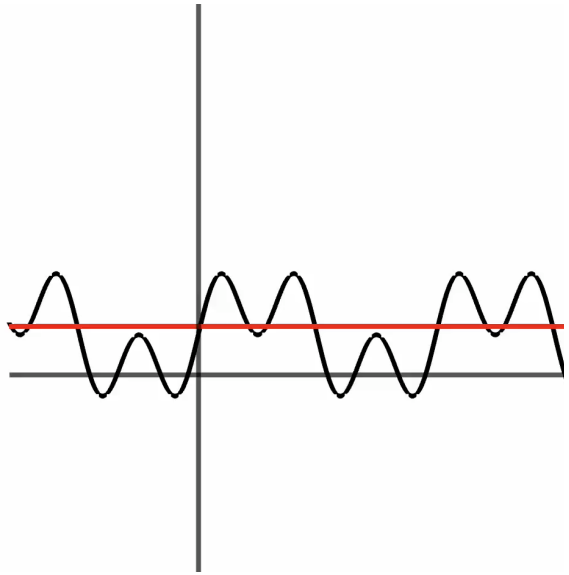
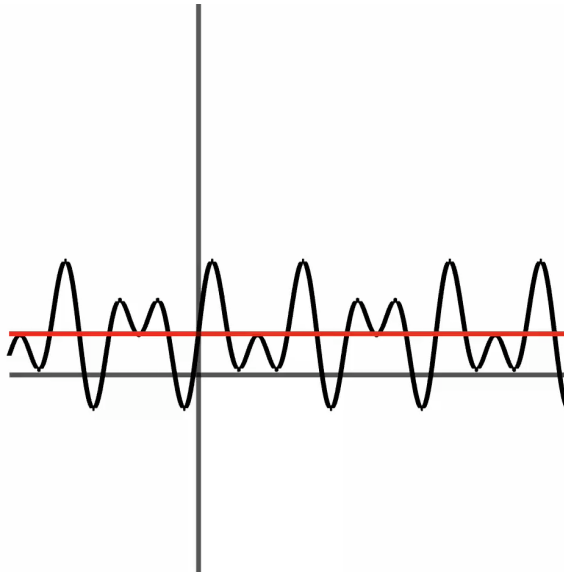
$$y = \sin(ax) + \sin(bx) + 1$$

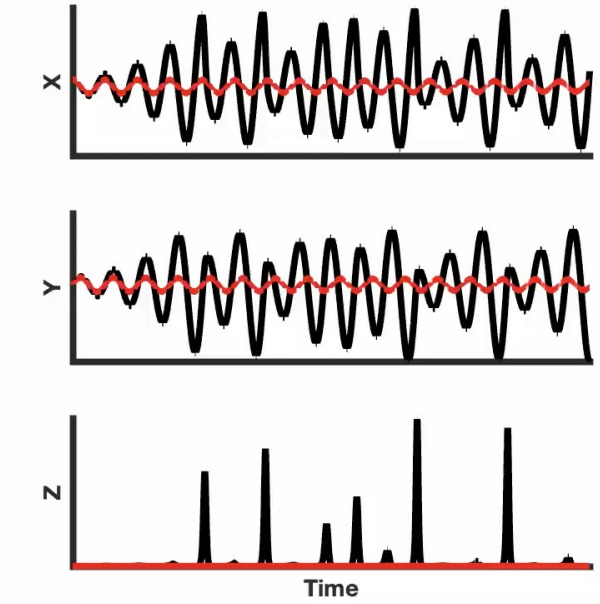


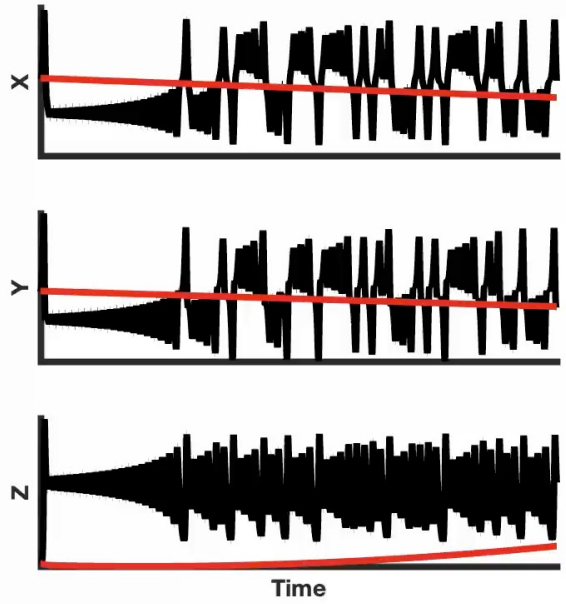


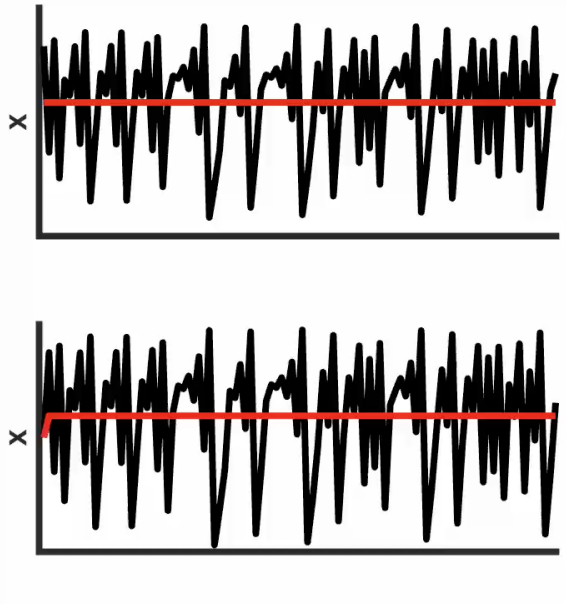
$$y = \sin(ax) + \sin(bx) + 1$$

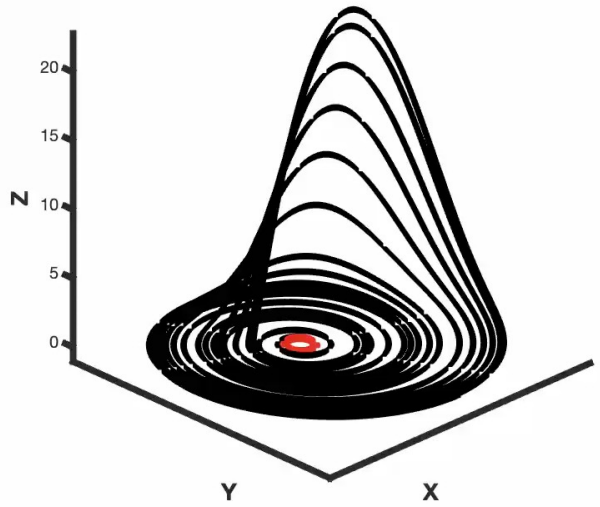


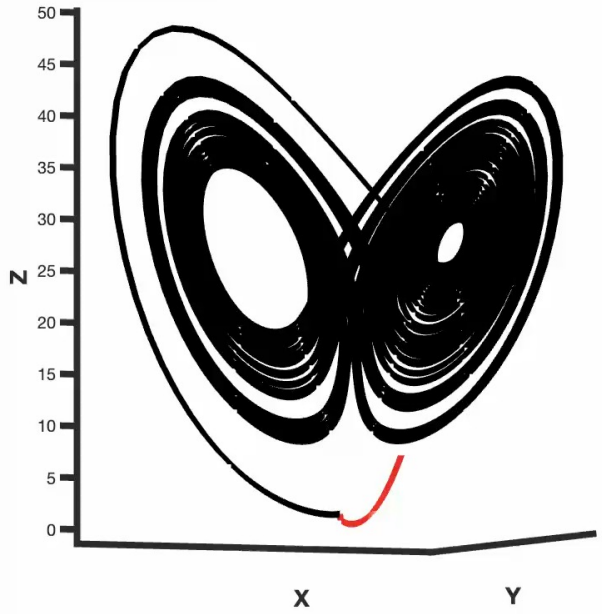







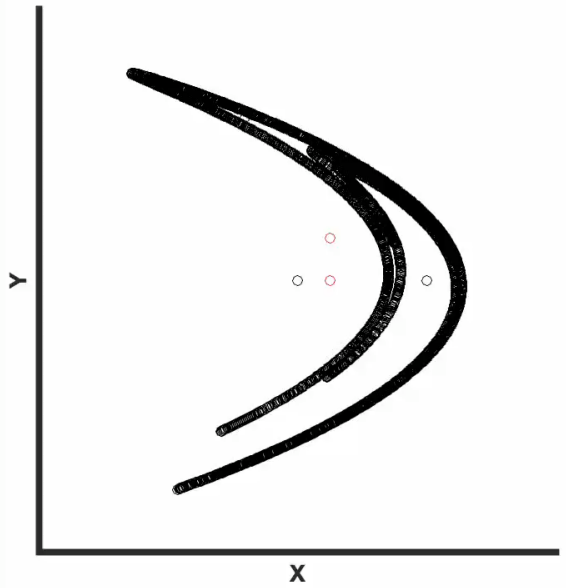

X
Y
Z
Time

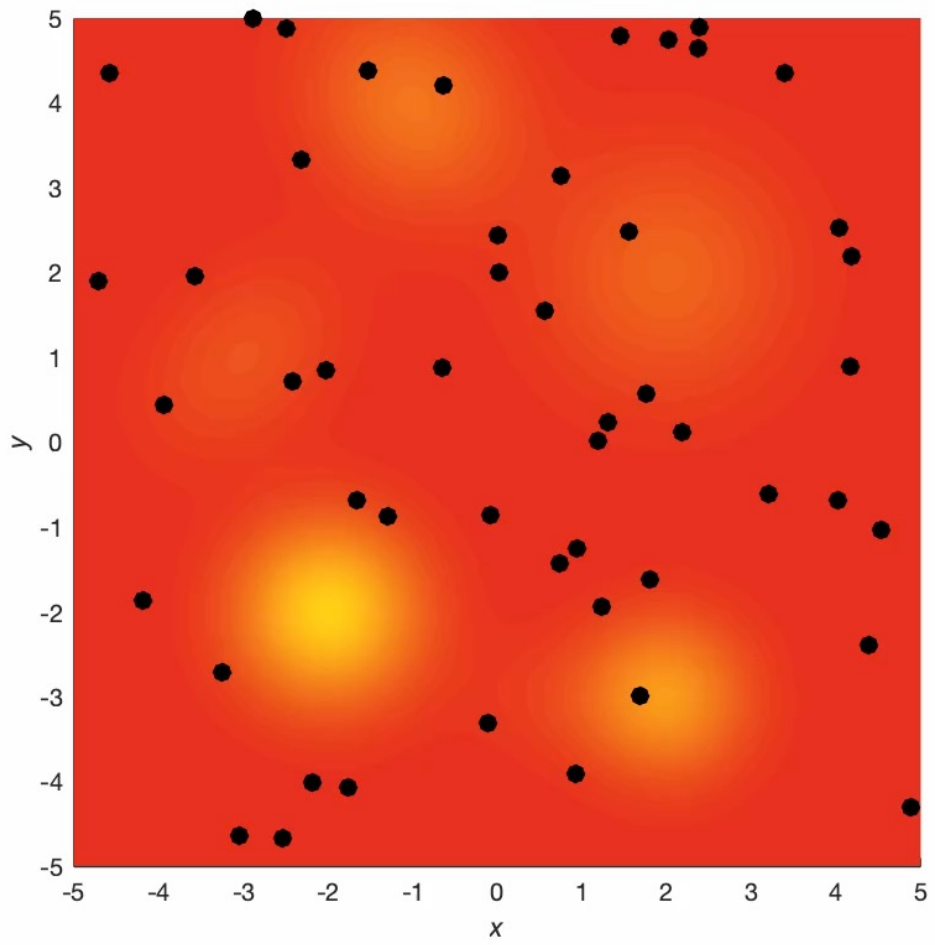

X
Y
Z
Time

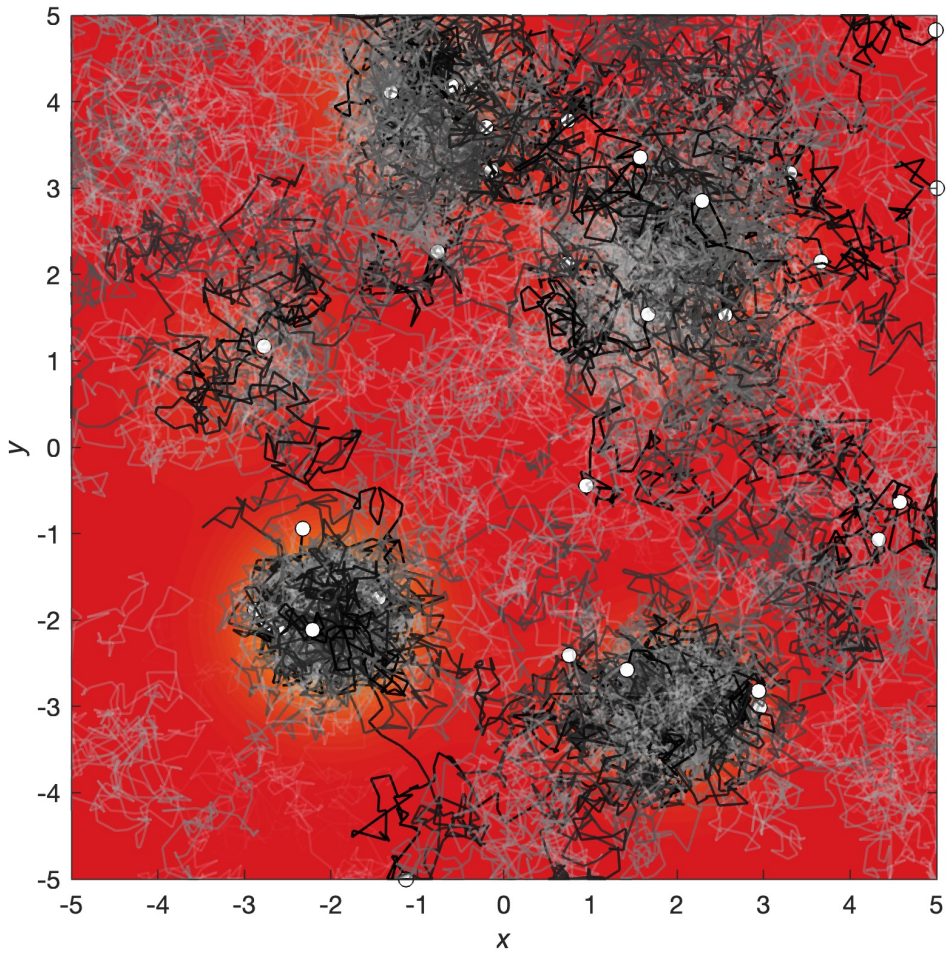

X
X

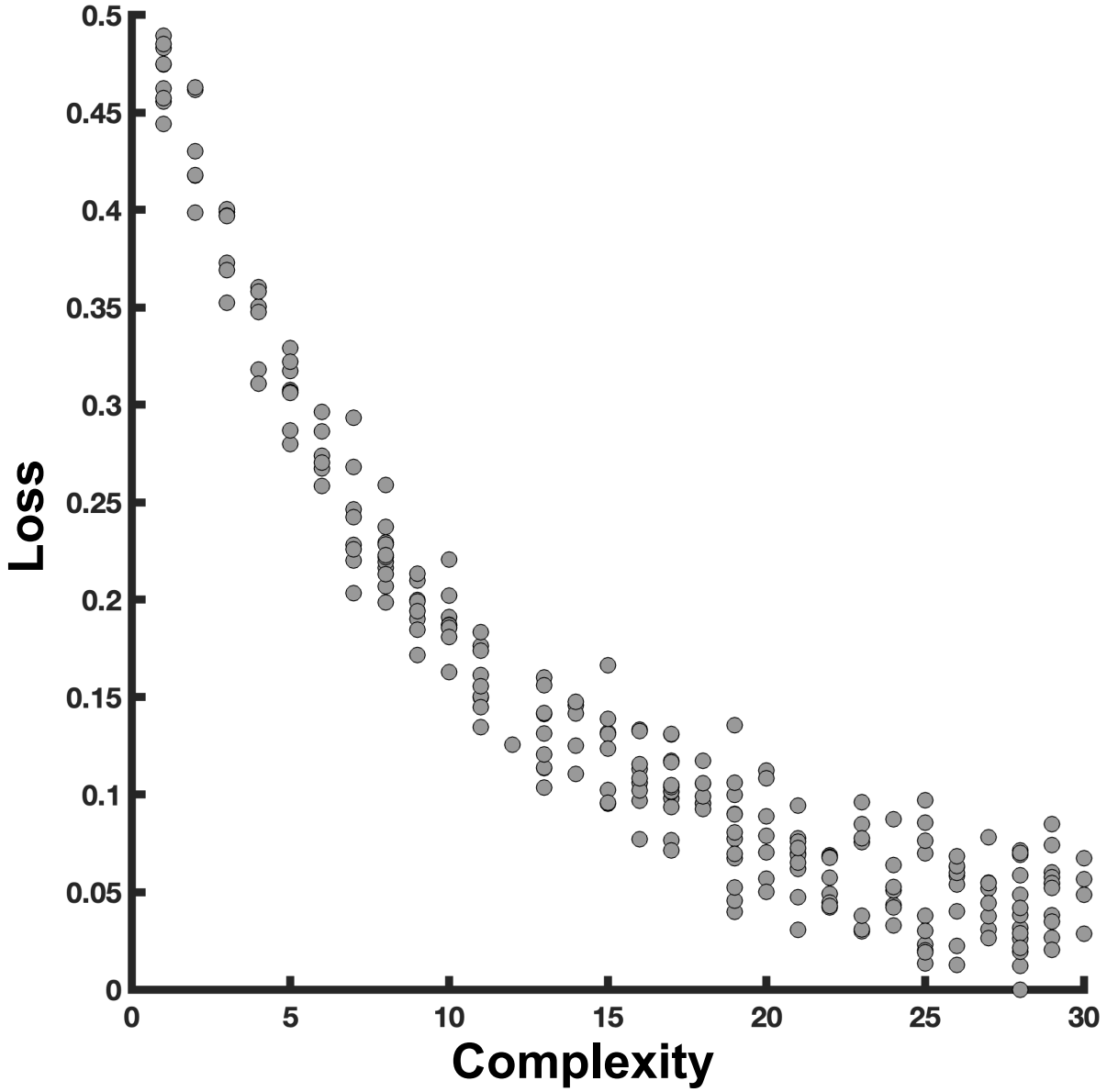

Z
20
15
10
5
0
Y
X


Z
50
45
40
35
30
25
20
15
10
5
0
X
Y


Y
X


5
4
3
2
1
0
-1
-2
-3
-4
-5
y
-5
-4
-3
-2
-1
0
1
2
3
4
5
x


5
4
3
2
1
0
-1
-2
-3
-4
-5
y
-5
-4
-3
-2
-1
0
1
2
3
4
5
x

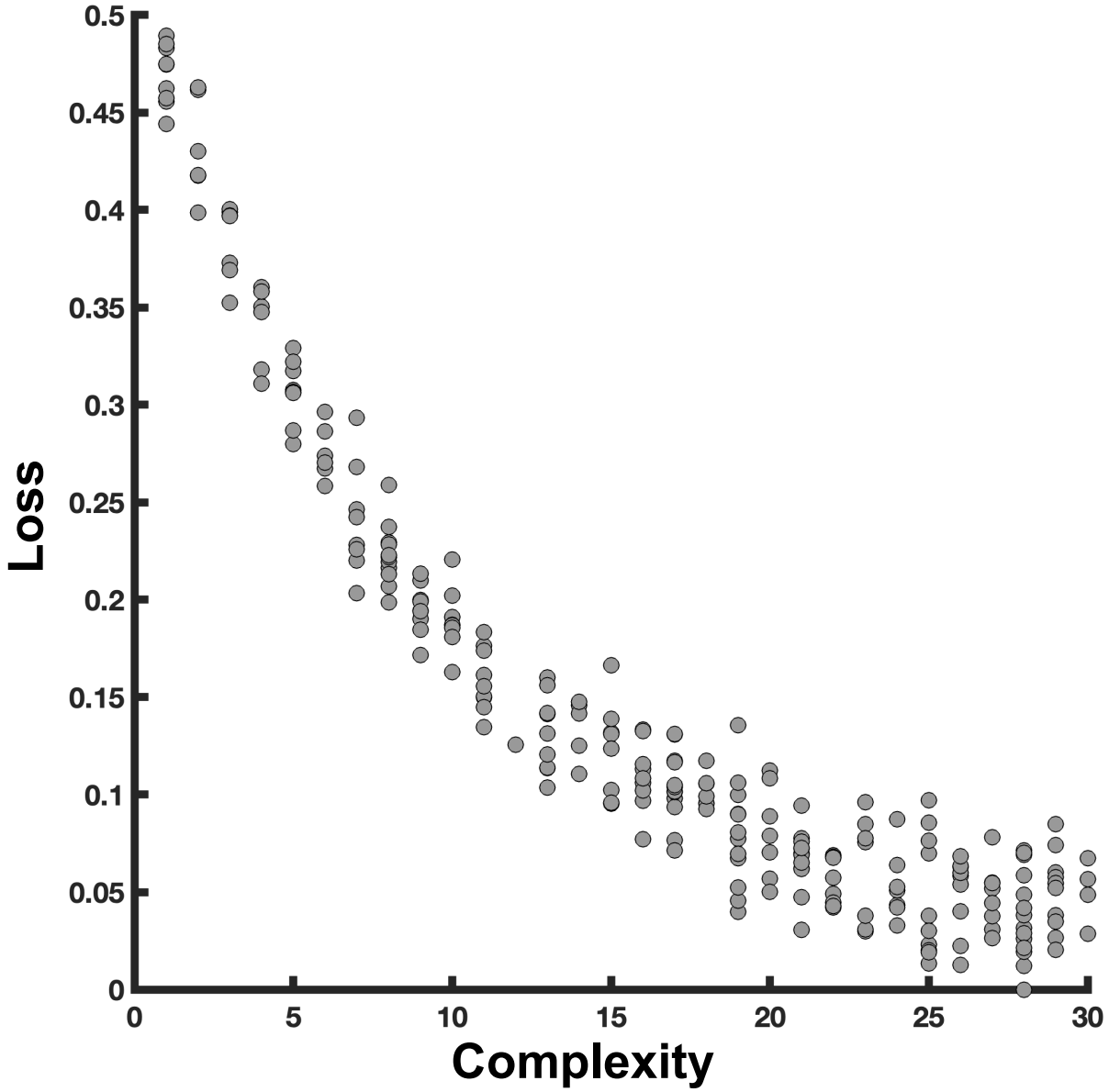

Loss
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Complexity

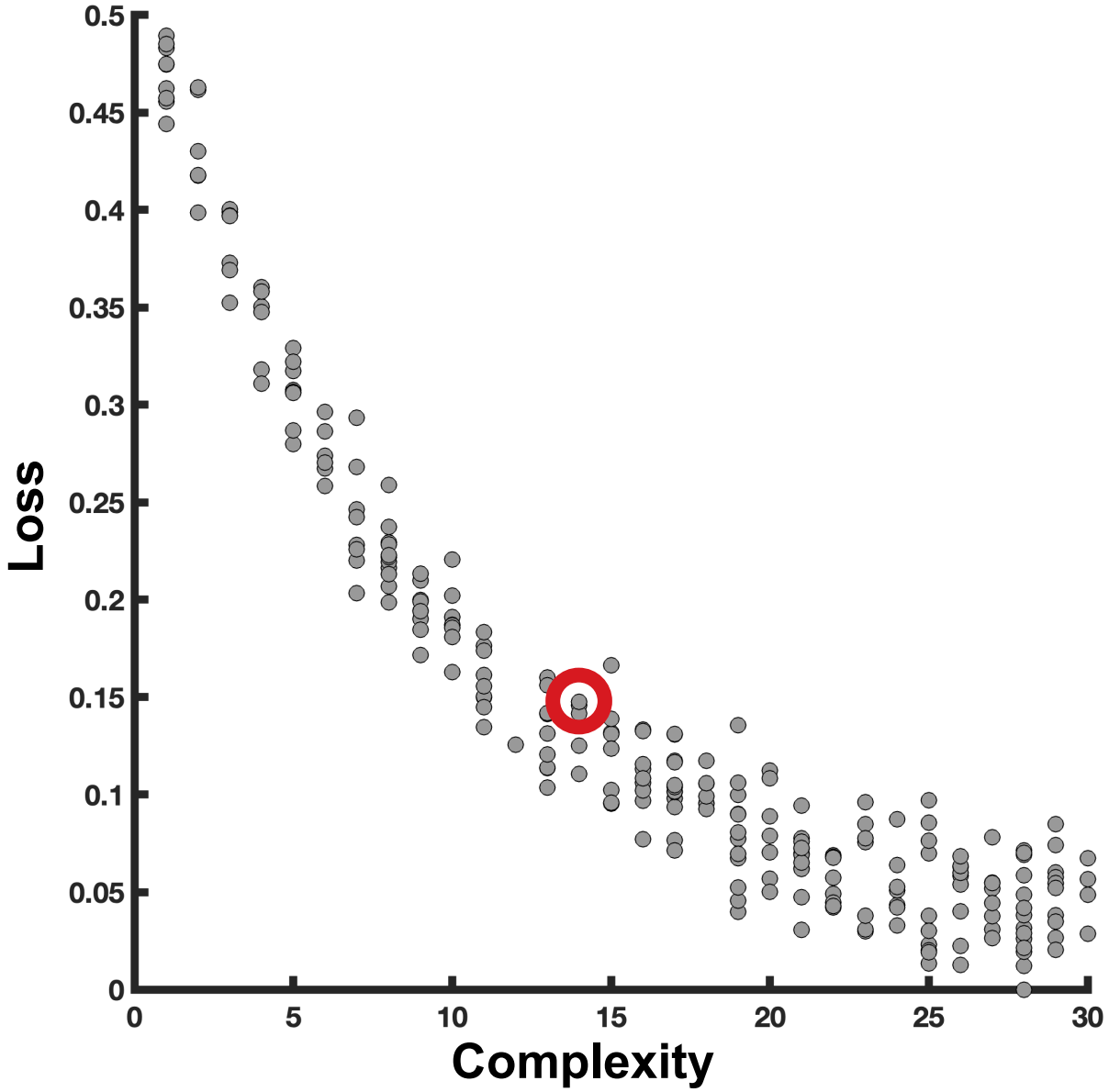
What is the BEST model?

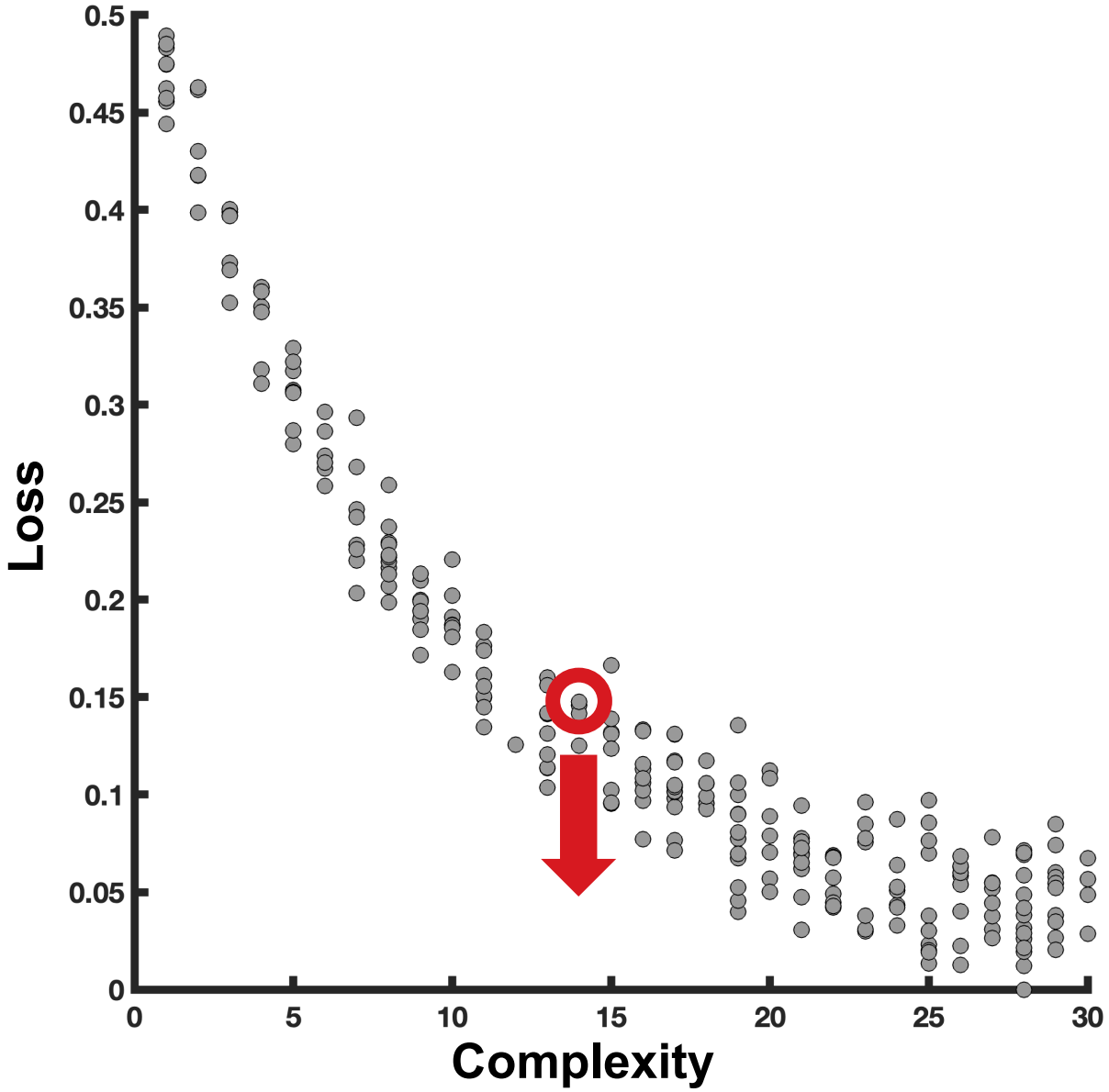
Accuracy? Complexity? Efficiency?

Pareto Front

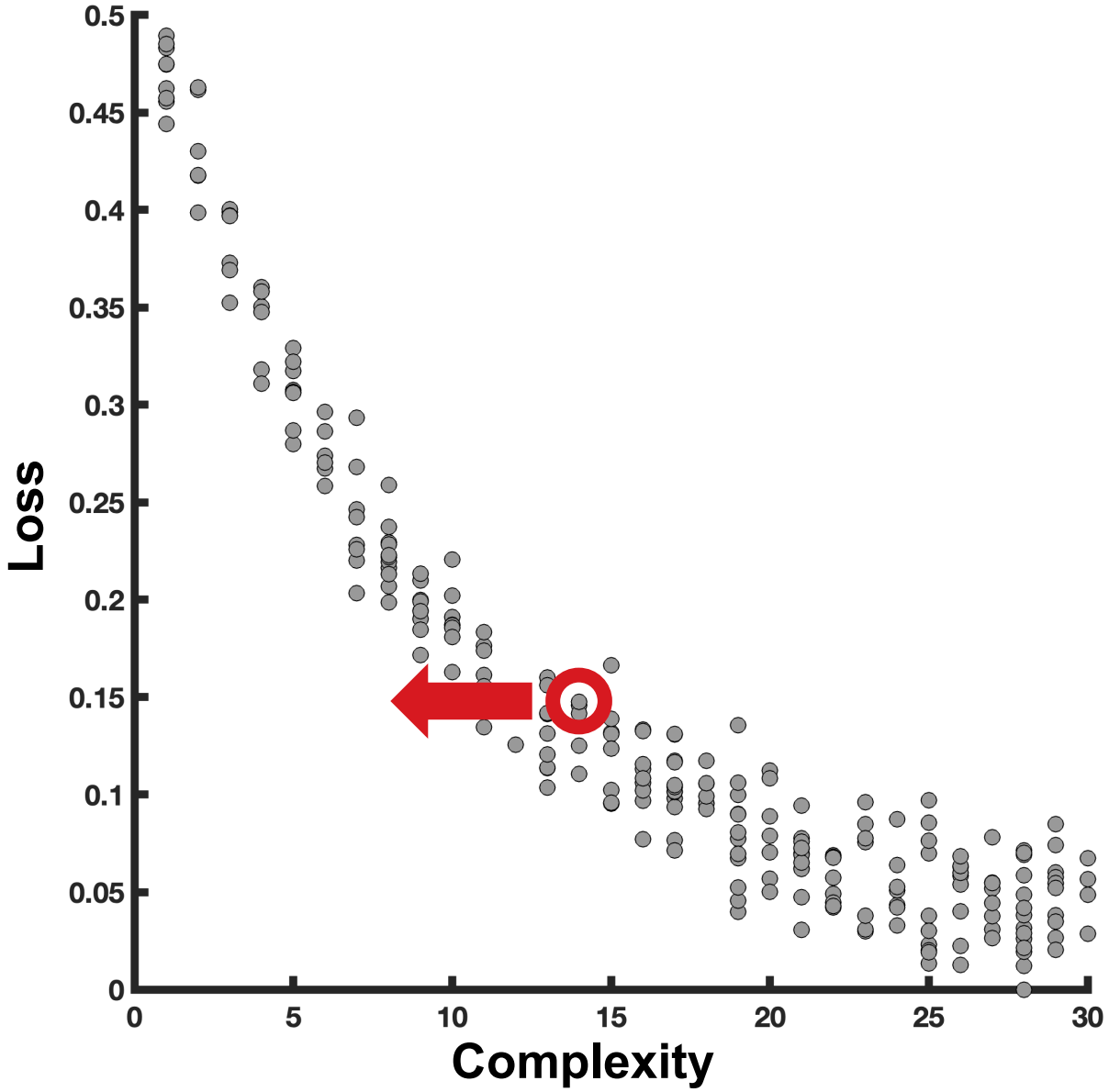
In multi-objective optimization, a Pareto front (or Pareto frontier) is the set of all non-dominated solutions—points for which every objective function cannot improve simultaneously.


Loss
Complexity
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30

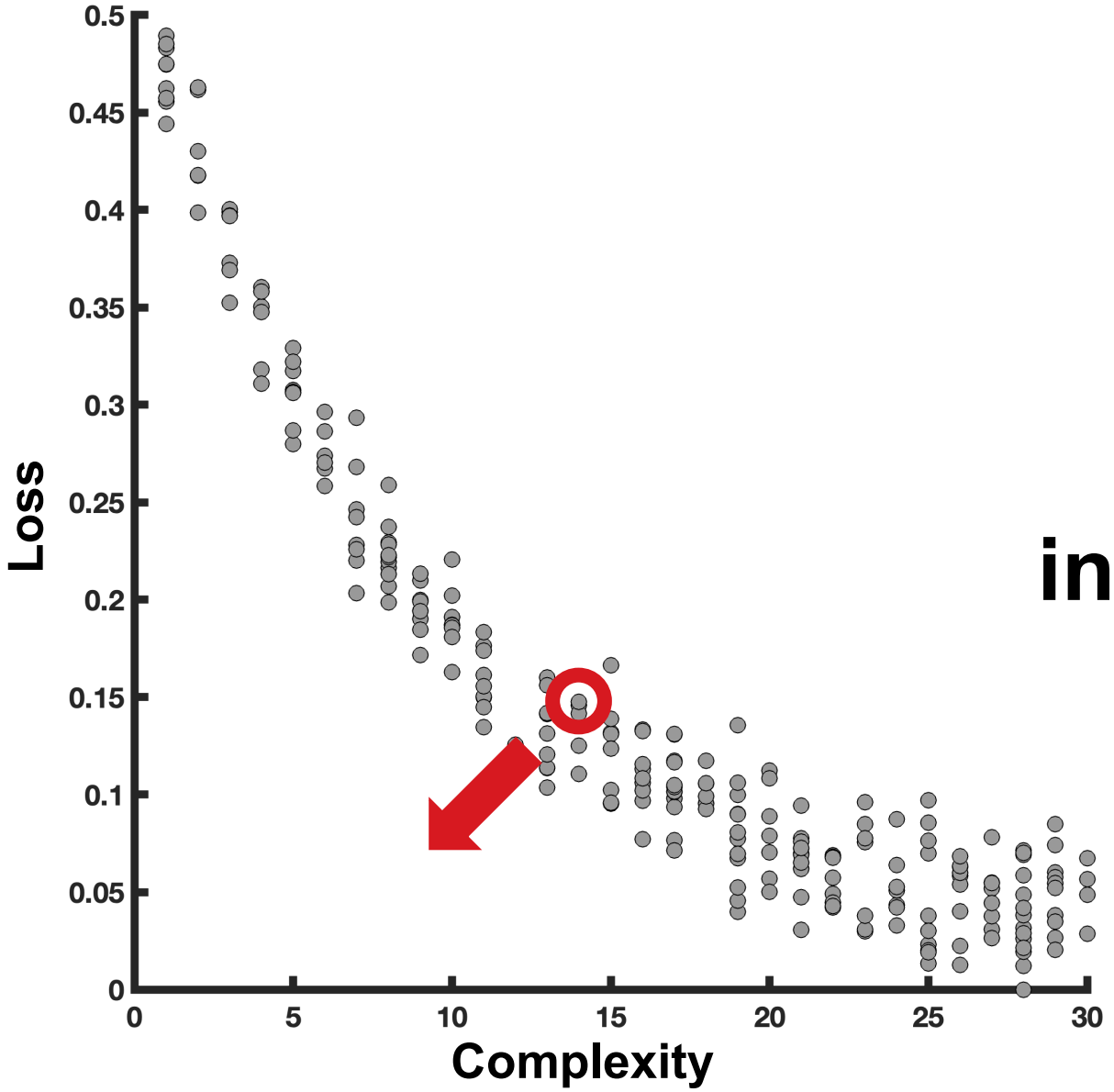

Loss
Complexity
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30



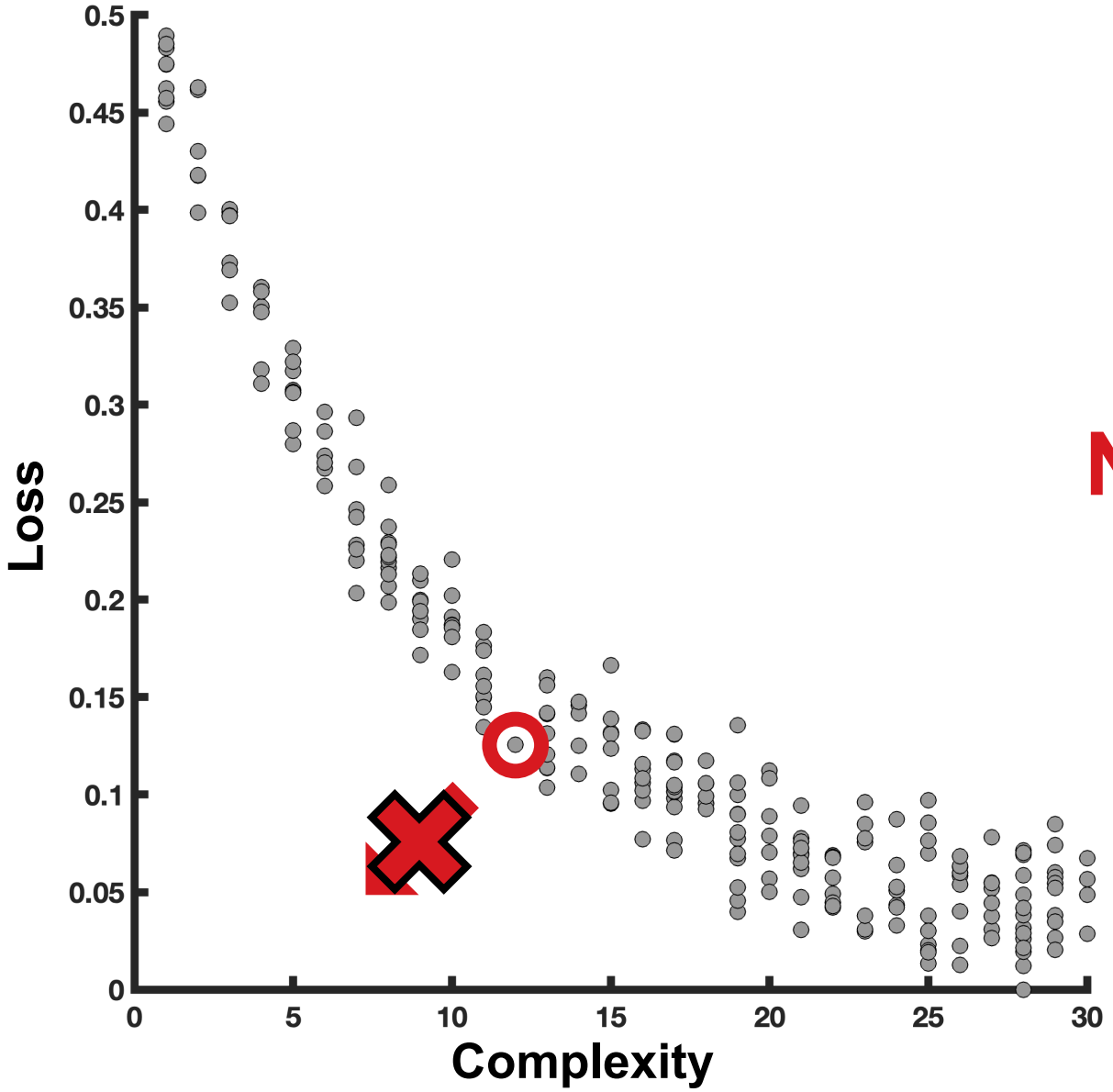
Increased Accuracy



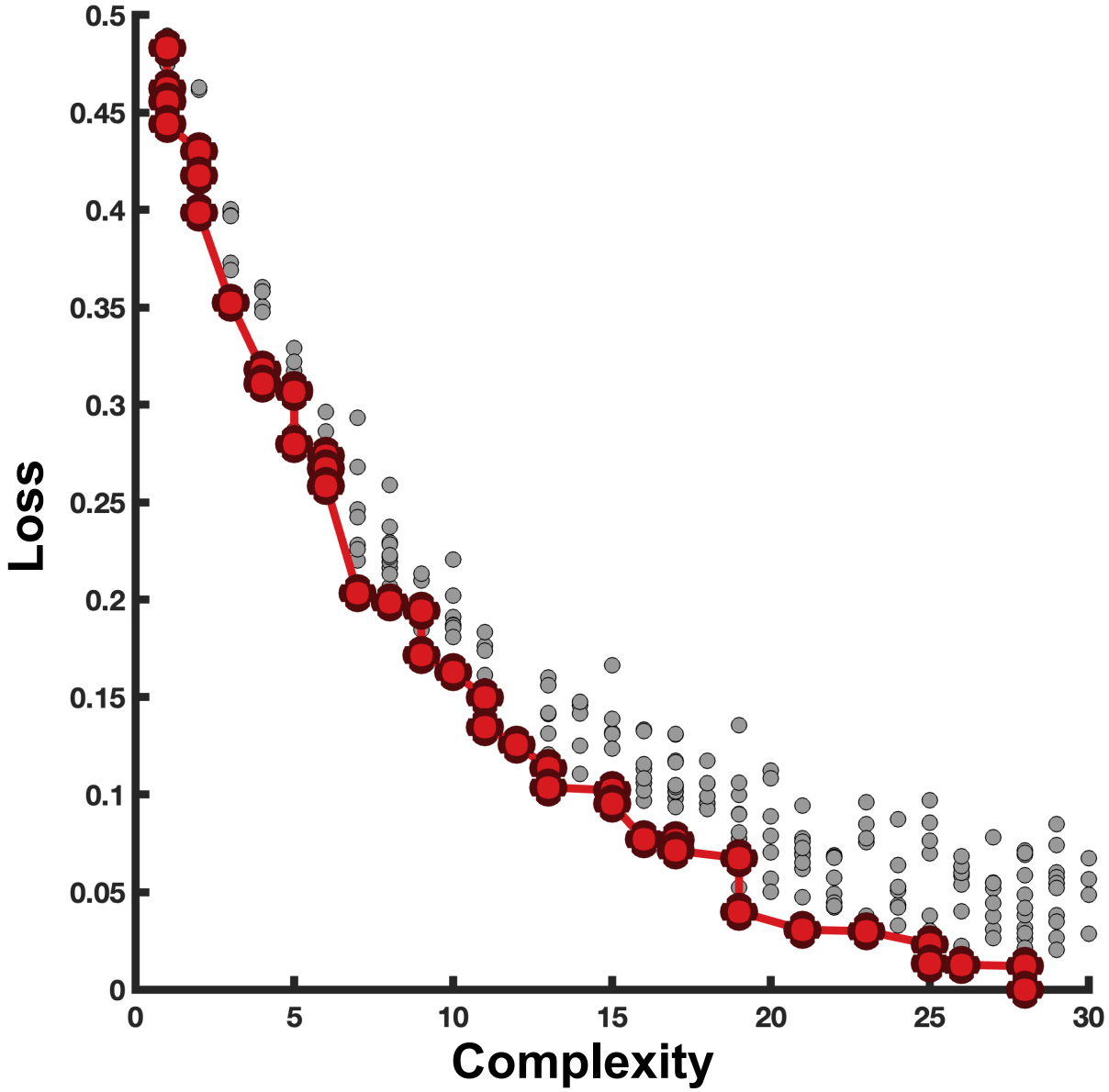
Increased Flexibility



Simultaneous Increase in Accuracy and Flexibility



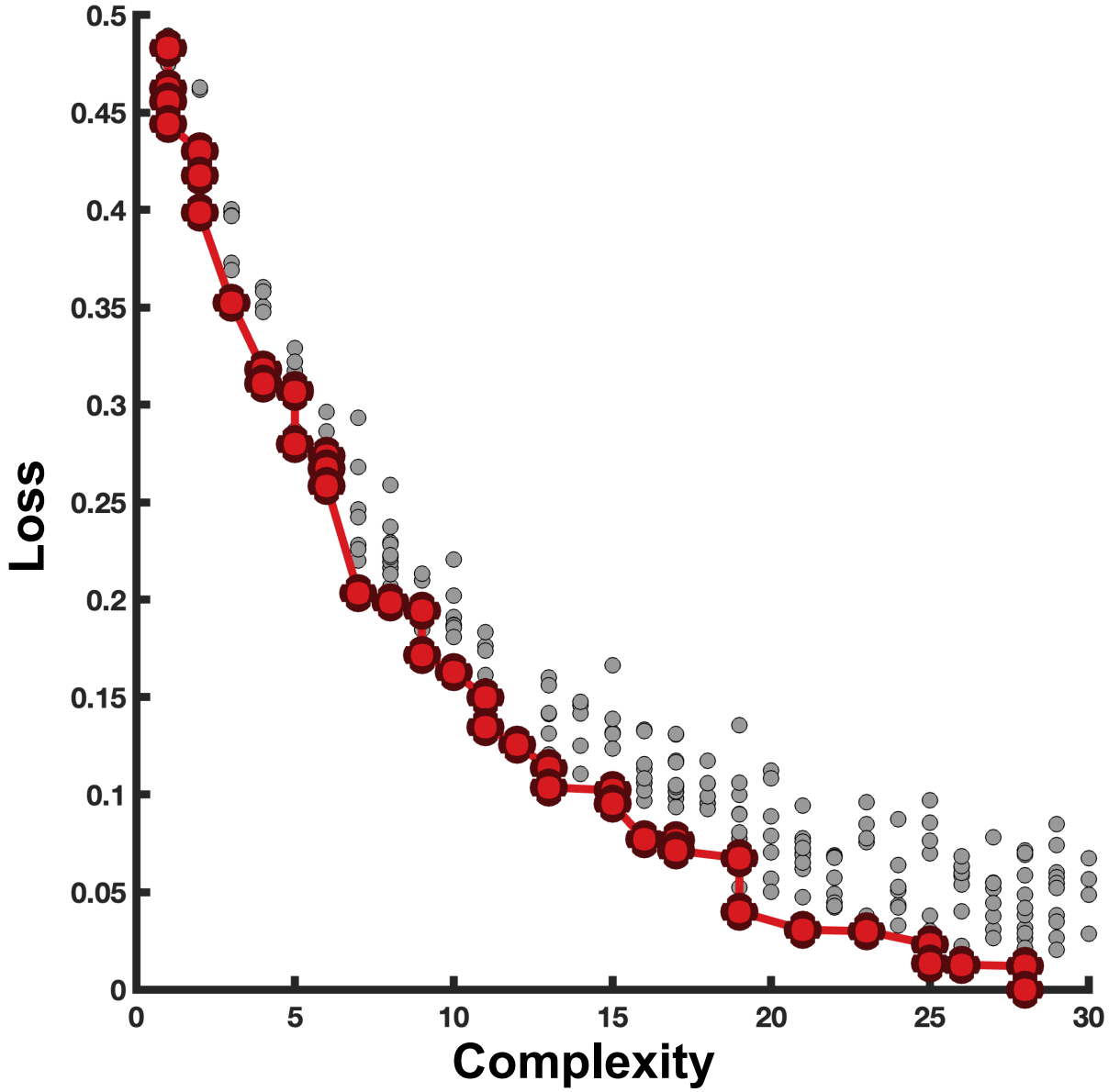
Non-dominated solution



Pareto Front

DEMO TIME!

Best Practices


Loss
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
Complexity

Model Selection

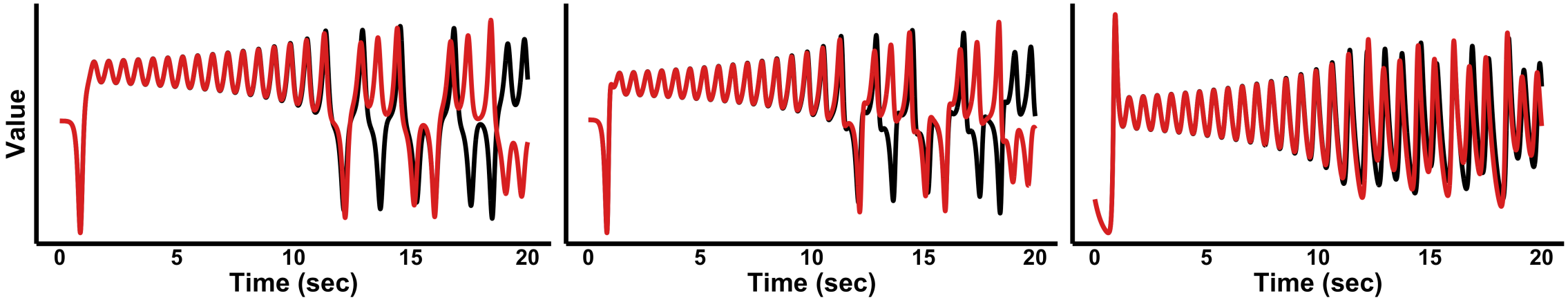
Multi-objective: Balance between accuracy and flexibility

Different Approaches of Model Selection

- Statistics-based
- Geometrics-based
- Semantics-based
- Miscellaneous

Computational Sensitivity

$$\begin{cases} dx/dt = 10(y - x) \\ dy/dt = x(28 - z) - y \\ dz/dt = xy - 8/3z \end{cases} \qquad \begin{cases} d\,x/d\,t = 9.996(y - x) \\ d\,y/d\,t = x(28.000 - z) - y \\ d\,z/d\,t = xy - 2.666z \end{cases}$$



Computational Sensitivity

1. Round coefficients and constants to closest integers

- 0.001 ➔ 0
- 0.999 ➔ 1
- 2.001 ➔ 2

2. Substitute coefficients and constants with closest, yet meaningful, rational and / or irrational numbers

- 0.333 ➔ $\frac{1}{3}$
- 2.718 ➔ e
- 3.142 ➔ π

Computational Sensitivity

3. Include context-specific coefficients and constants as variables

- Periodic data ➔ π
- Gravitation force ➔ g
- Secrets of the universe ➔ e

Probabilistic Nature

Cons ("*Single, reliable equation!*")

Mitigating variance

- Cross-validation
- Ensemble methods
- Multi-run consensus
- Better algorithms
  - Stability measures can help assess variance of SR algorithms
- Parameter tuning: larger population size, longer run time, etc.

May incur more bias

Probabilistic Nature

Pros ("*Rich information about the data!*")

Leveraging variance

- Distribution of variables and structures
- Allow informed choice
- May reflect non-unique solutions

Probabilistic Nature

Pros ("*Rich information about the data!*")

Leveraging variance

- Distribution of variables and structures
- Allow informed choice
- May reflect non-unique solutions

Run multiple bouts of symbolic regression!

Model Interpretation

Model Interpretation

Discovered models are NOT ‘TRUTH’

Instead, they are CANDIDATE explanations

Model Interpretation

Discovered models are NOT ‘TRUTH’

Instead, they are CANDIDATE explanations

Do the terms make biological and physical sense?

Model Interpretation

Discovered models are NOT ‘TRUTH’

Instead, they are CANDIDATE explanations

Do the terms make biological and physical sense?

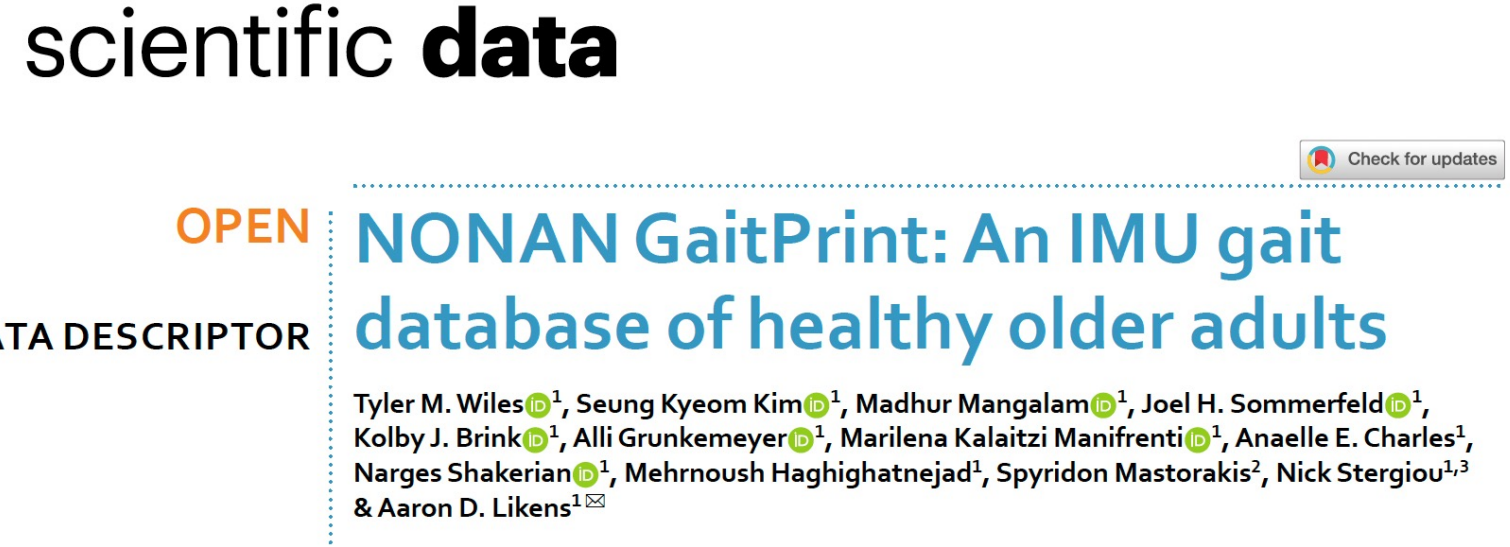
Which terms correspond to known mechanisms?

Complete Example

Aim

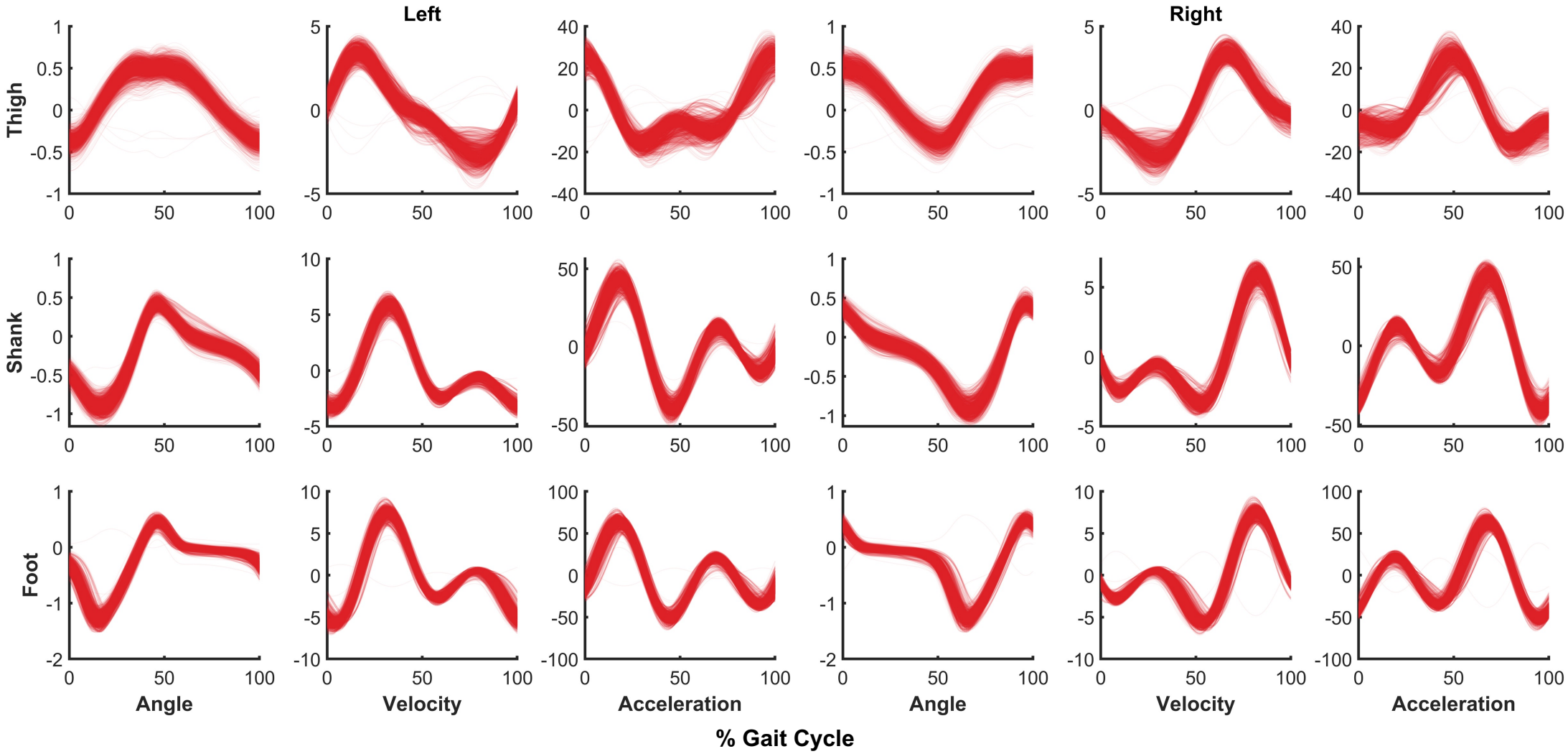
1. Discover the equations of motion of gait

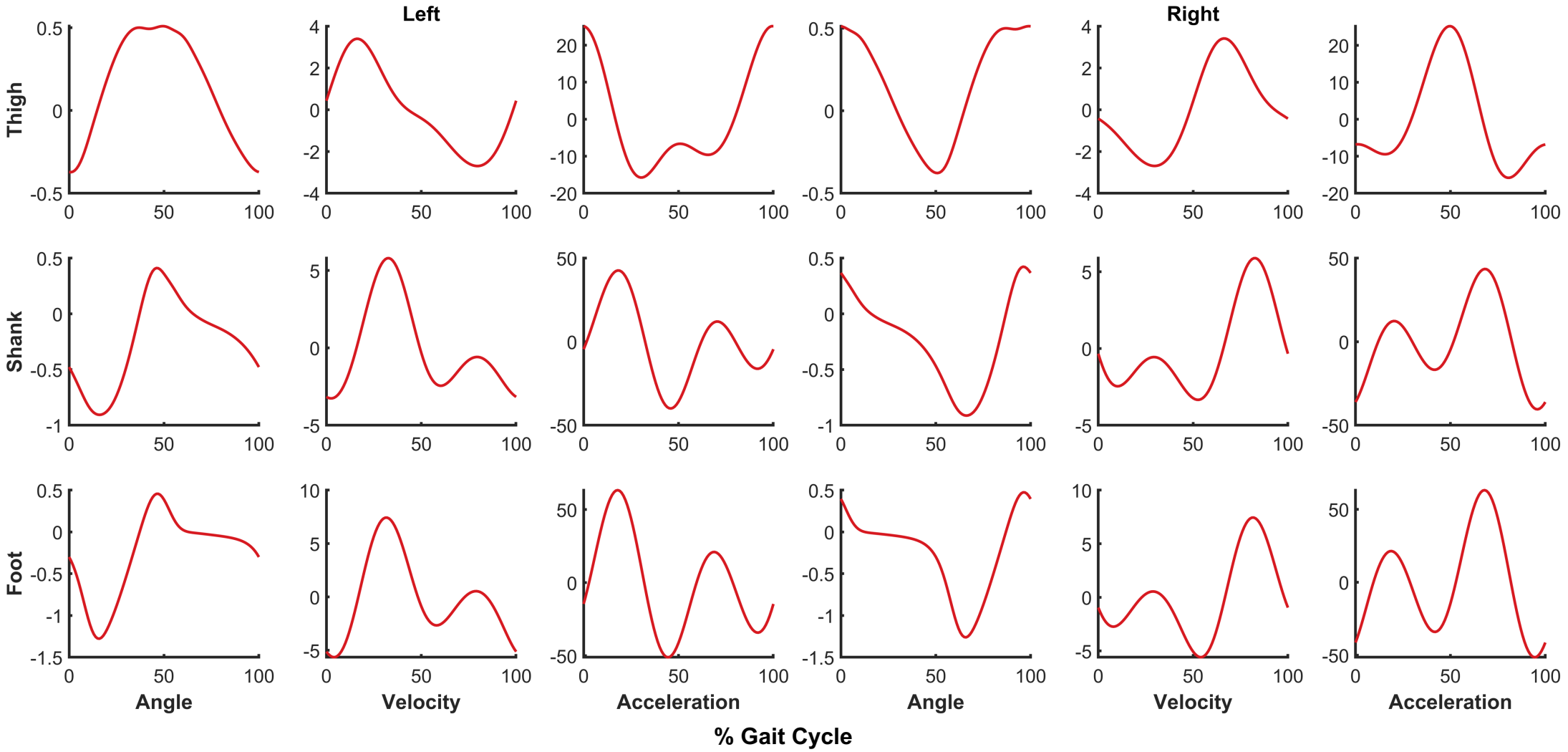
2. Predict segment angular acceleration



Data Description

- Healthy young, middle, older adults (N = 35 / 36 / 41)
- Overground walking trials
  - 4 minutes/trial
  - 9 trials/session
  - 2 sessions one week apart
- Inertial measurement units (200Hz)
- Total of 1,944 trials


Left
Right
Thigh
Shank
Foot
Angle
Velocity
Acceleration
Angle
Velocity
Acceleration
% Gait Cycle


Left
Right
Thigh
Shank
Foot
Angle
Velocity
Acceleration
Angle
Velocity
Acceleration
% Gait Cycle

Symbolic Regression

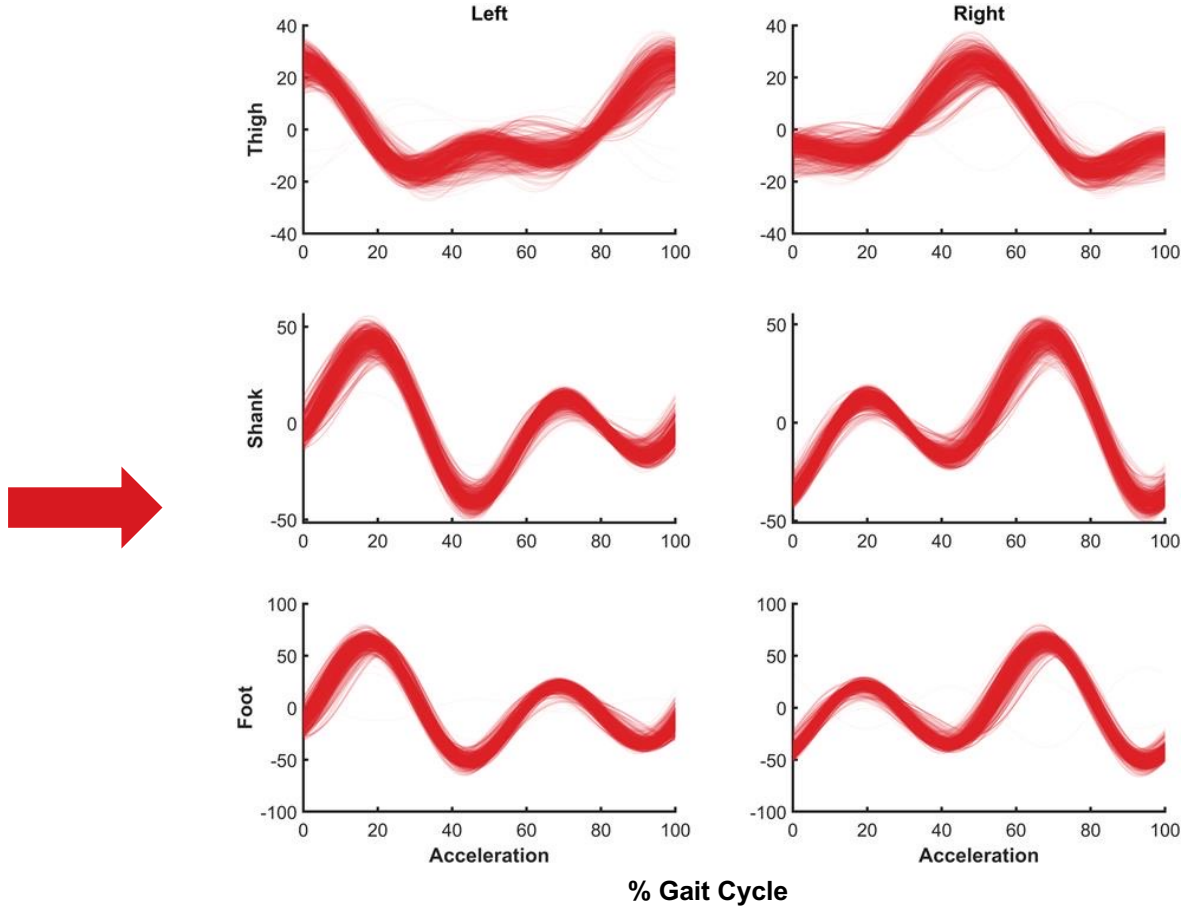
- **Outcome Variables**
  - **Thigh, Shank, Foot Pitch Angular Acceleration ($\ddot{\theta}_{thigh}, \ddot{\theta}_{shank}, \ddot{\theta}_{foot}$)**
- **Input Variables**
  - **Bilateral Thigh, Shank, Foot Pitch Angle ($\theta_{thigh}, \theta_{shank}, \theta_{foot}$)**
  - **Bilateral Thigh, Shank, Foot Pitch Angular Velocity ($\dot{\theta}_{thigh}, \dot{\theta}_{shank}, \dot{\theta}_{foot}$)**
- **Operators: $+, -, \times, \div$**
- **Functions: sin, cos (with no nested functions)**

Symbolic Regression

- **Parameters**
  - **Max Equation Size: 50**
  - **Number of Iterations: 10,000**
  - **Number of Populations: 31 (Default)**
  - **Population Size: 27 (Default)**
  - **Number of Cycles per Iteration: 380 (Default)**
  - **Loss Function: L1 Norm**

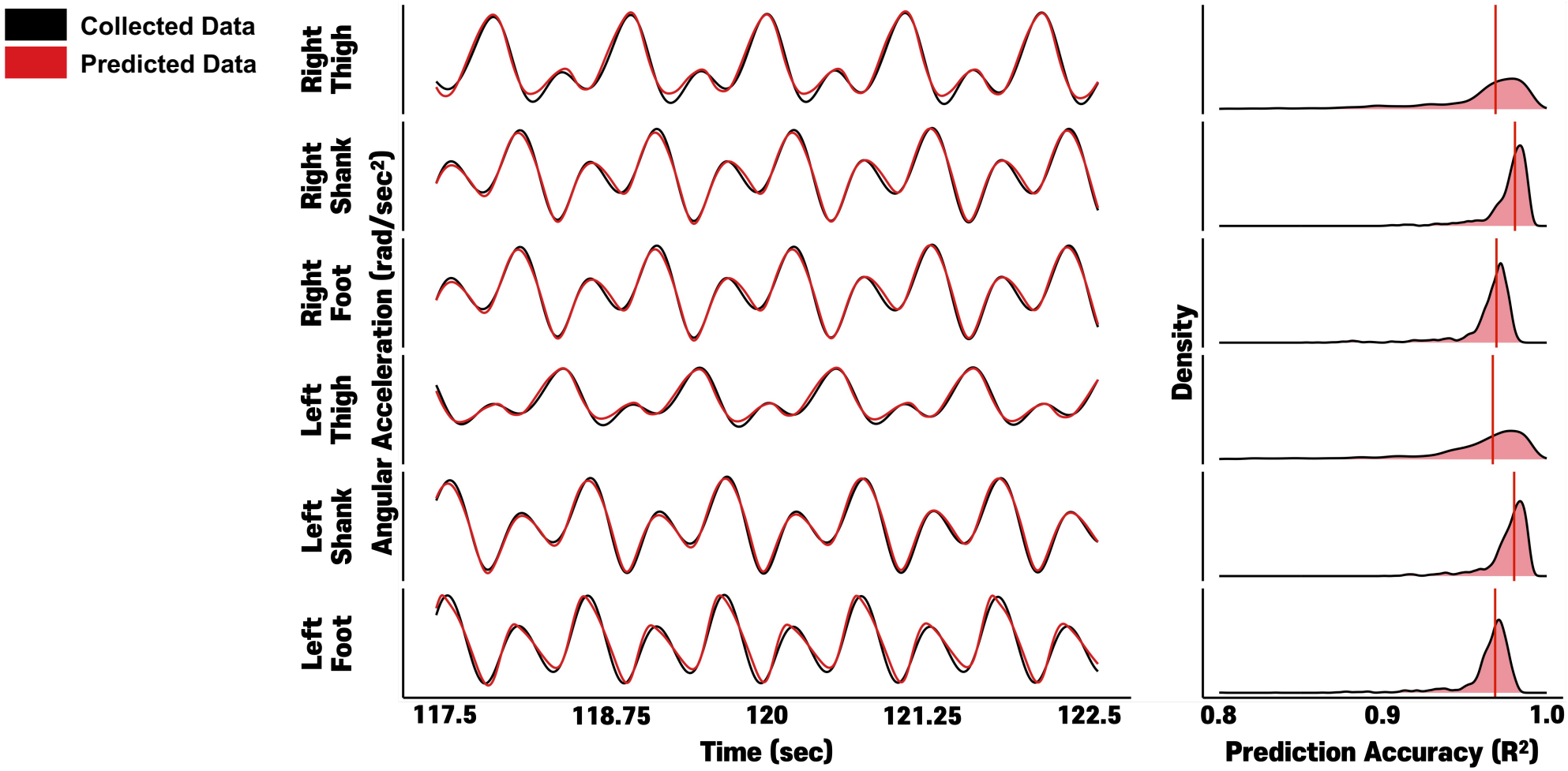
$$\begin{cases} \ddot{\theta}_{thigh} = \theta_{ipsi.thigh}\left(\alpha_{t1}\theta_{contra.thigh} + \alpha_{t2}\right) + \alpha_{t0} \\ \ddot{\theta}_{shank} = \alpha_{s1}\theta_{ipsi.shank} + \alpha_{s2}\theta_{contra.shank} + \alpha_{s0} \\ \ddot{\theta}_{foot} = \alpha_{f1}\theta_{ipsi.foot} + \alpha_{f2}\theta_{contra.foot} + \alpha_{f0} \end{cases}$$

$$\begin{cases} \ddot{\theta}_{thigh} = \theta_{ipsi.thigh}(\alpha_{t1}\theta_{contra.thigh} + \alpha_{t2}) + \alpha_{t0} \\ \ddot{\theta}_{shank} = \alpha_{s1}\theta_{ipsi.shank} + \alpha_{s2}\theta_{contra.shank} + \alpha_{s0} \\ \ddot{\theta}_{foot} = \alpha_{f1}\theta_{ipsi.foot} + \alpha_{f2}\theta_{contra.foot} + \alpha_{f0} \end{cases}$$



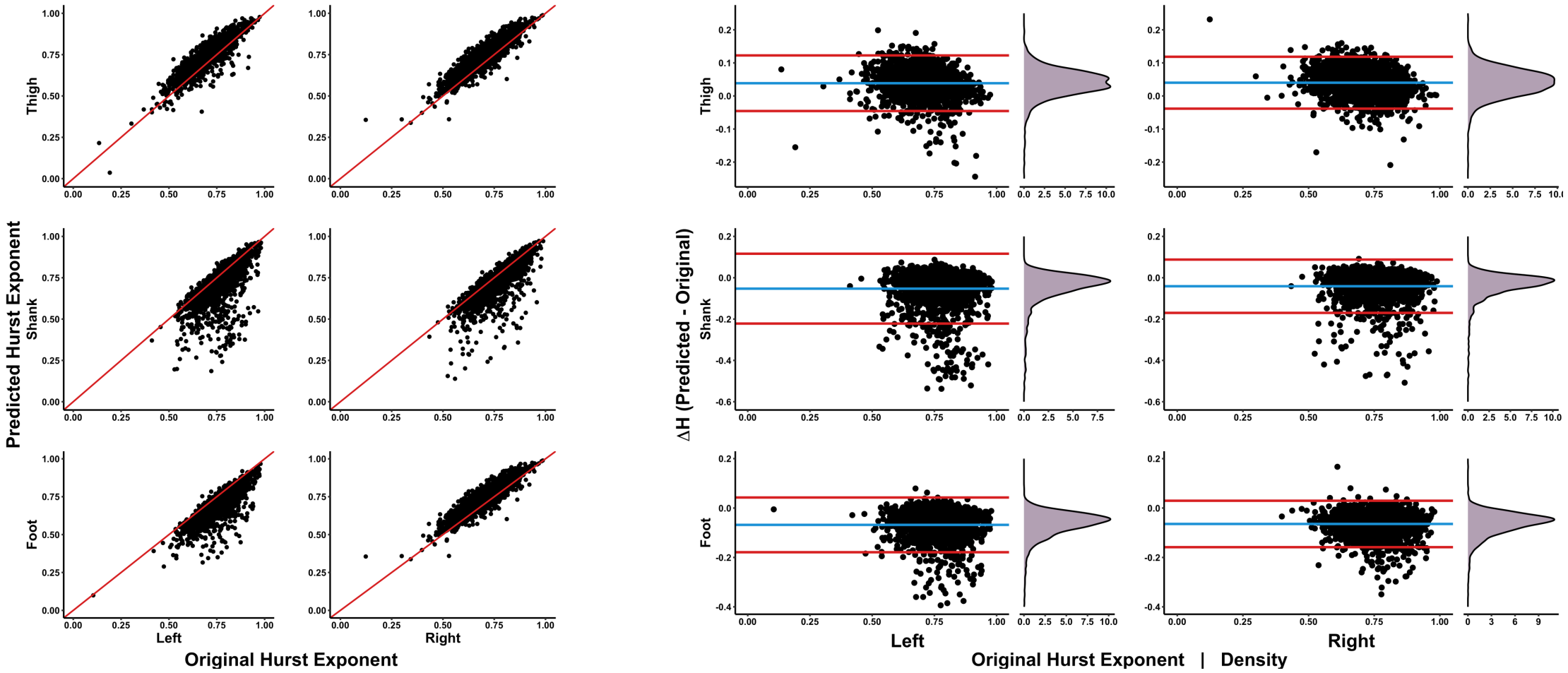
Angular Acceleration Prediction

Prediction Accuracy



Kim et al., 2025 (In Preparation)

Preservation of Gait Variability



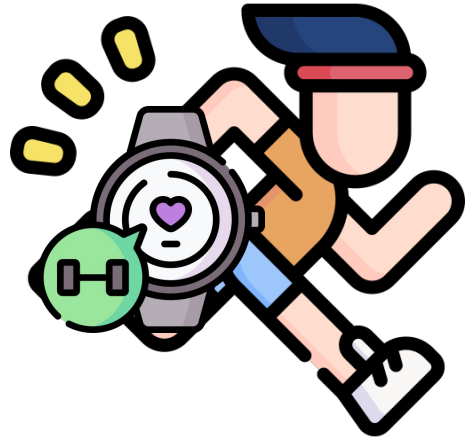
Kim et al., 2025 (In Preparation)

What’s Next?

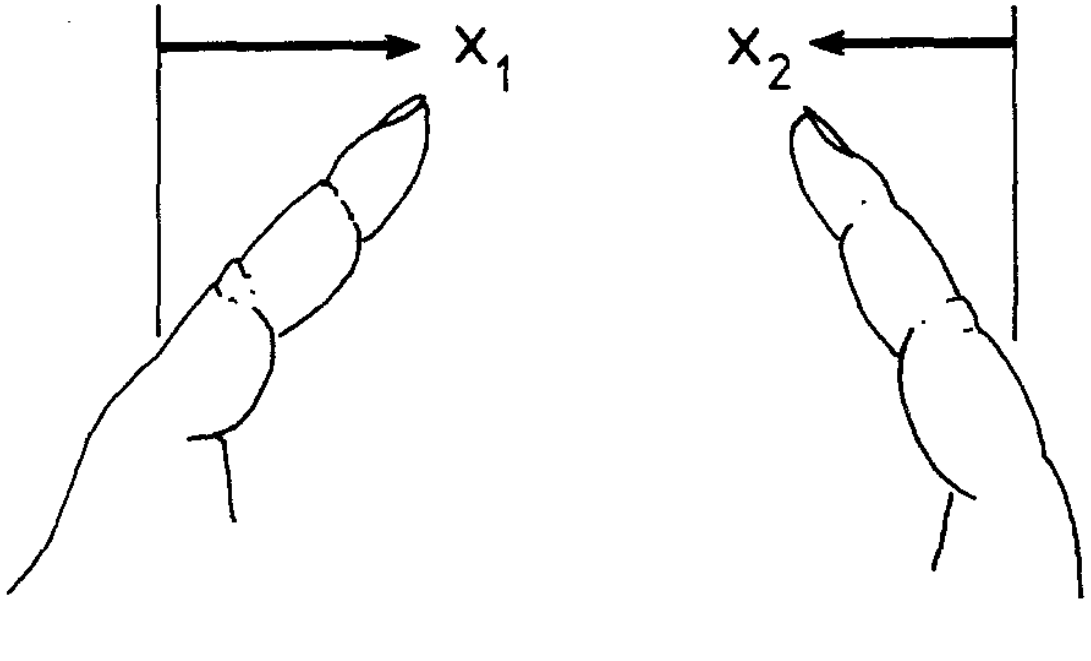
Discover

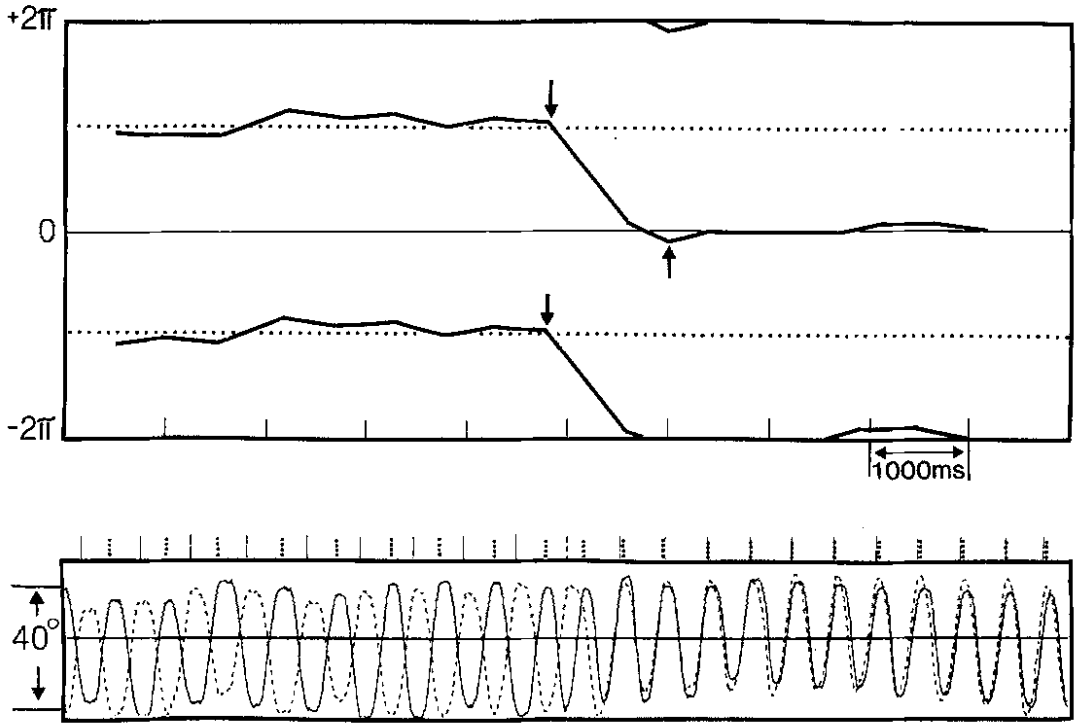
Interpret

Predict



Test

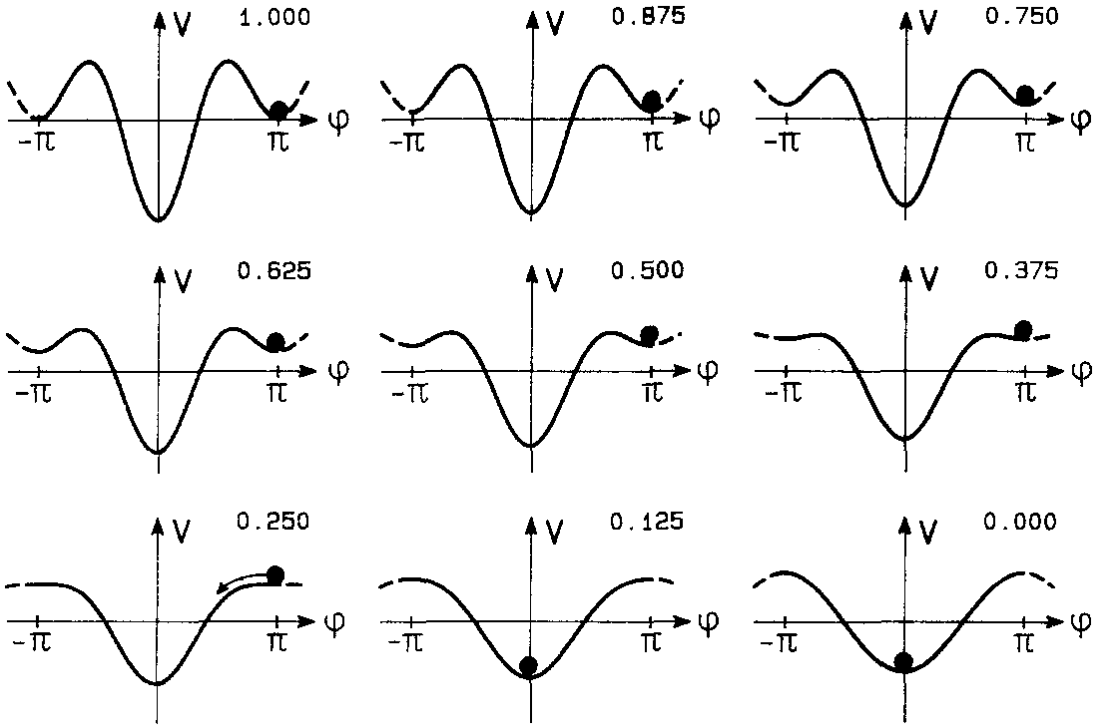

X_1
X_2

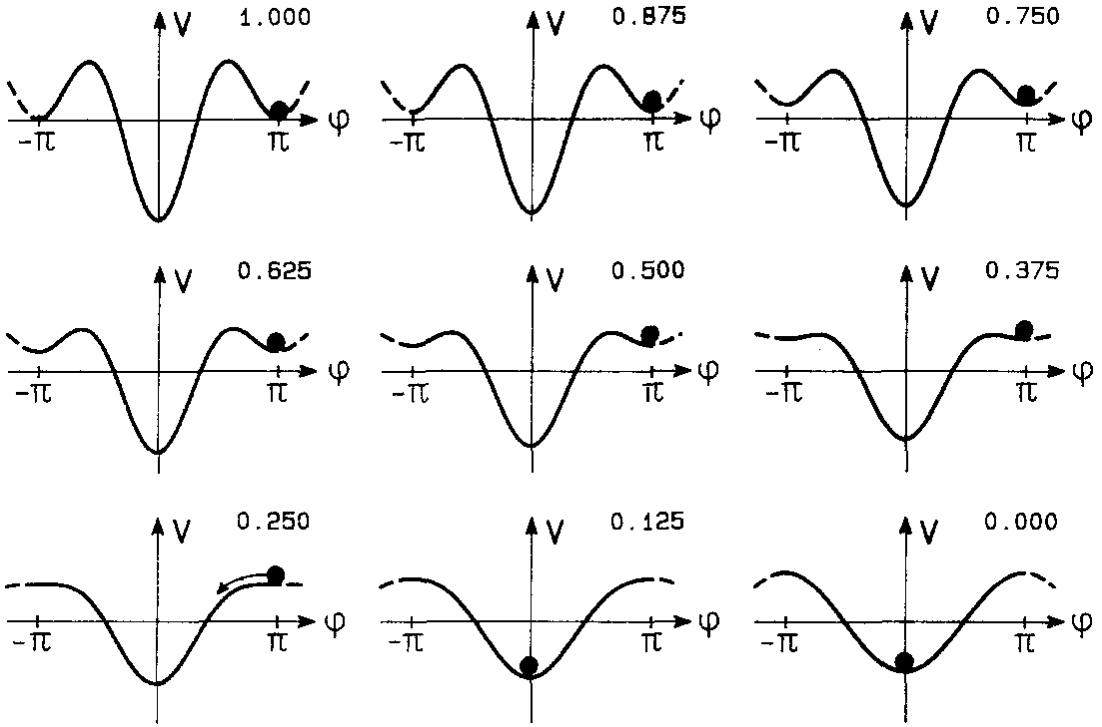

+2π
0
-2π
1000ms
40°

Haken-Kelso-Bunz (HKB) Model

$$V = -a\cos\phi - b\cos 2\phi$$

$$\dot{\phi} = a\sin\phi + 2b\sin 2\phi$$





Critical Slowing Down

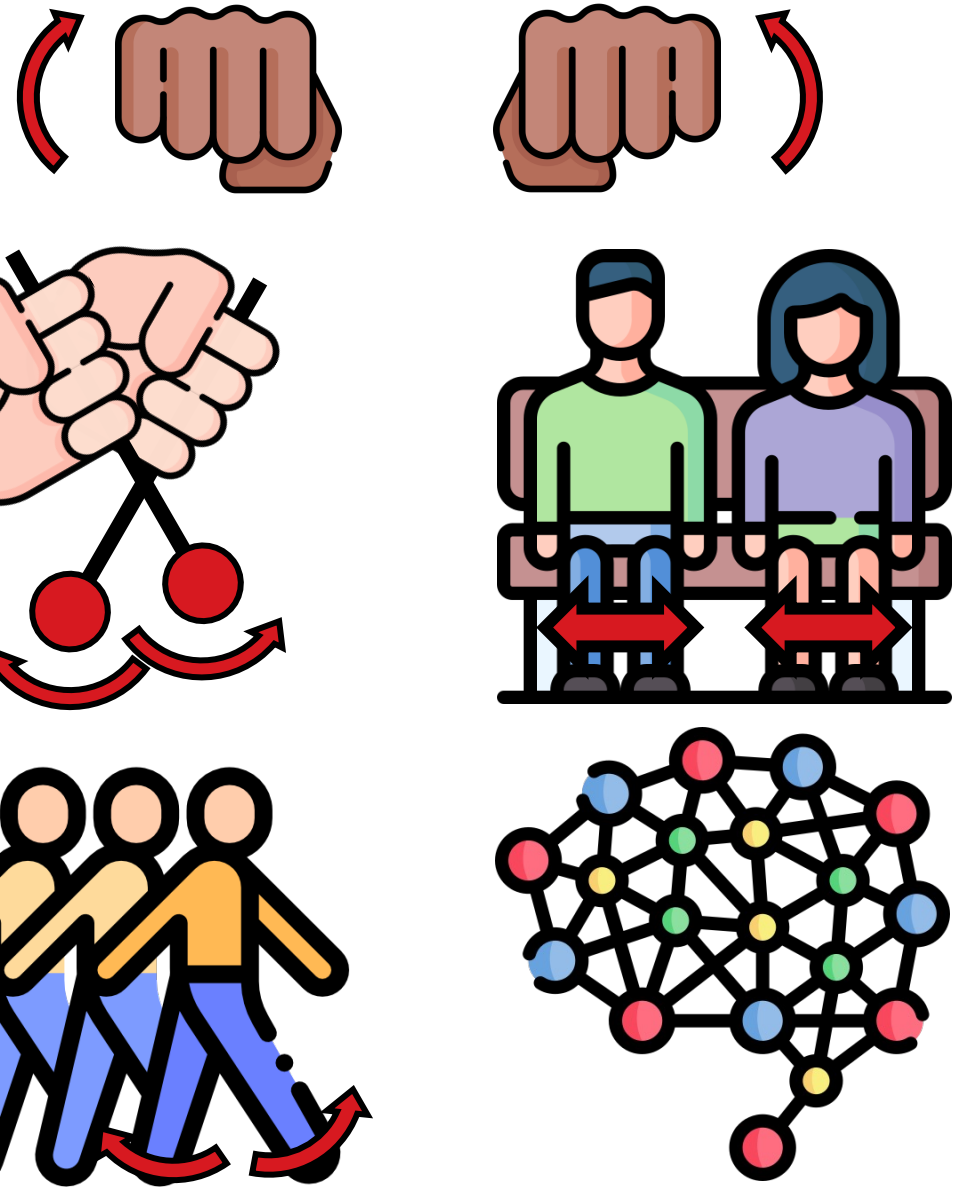
Critical Fluctuation

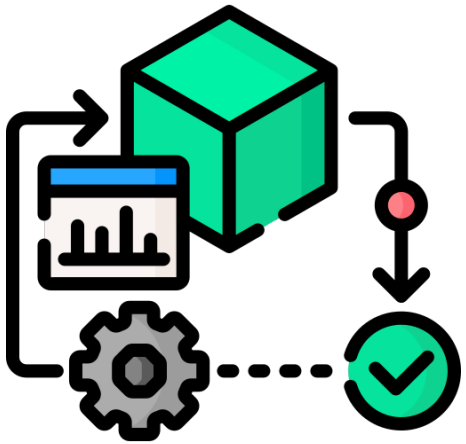
Hysteresis

Critical Slowing Down

Critical Fluctuation

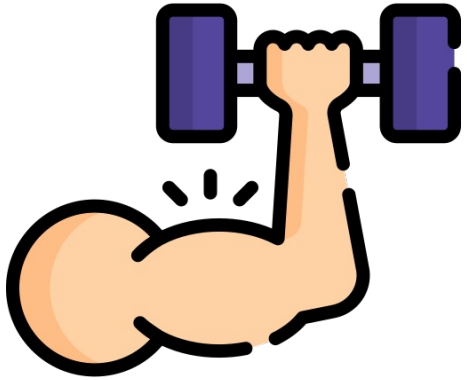
Hysteresis



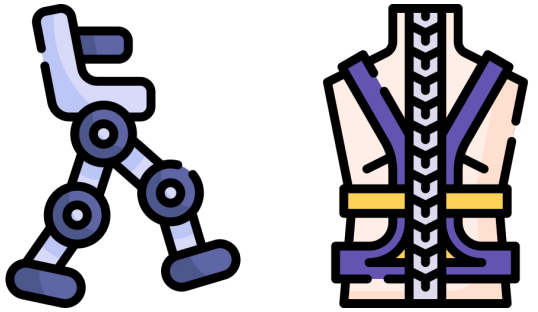


Rigorous Research Practice

Discovery of Biomechanical Law



Personalized Training Protocol



Human Computer Interaction