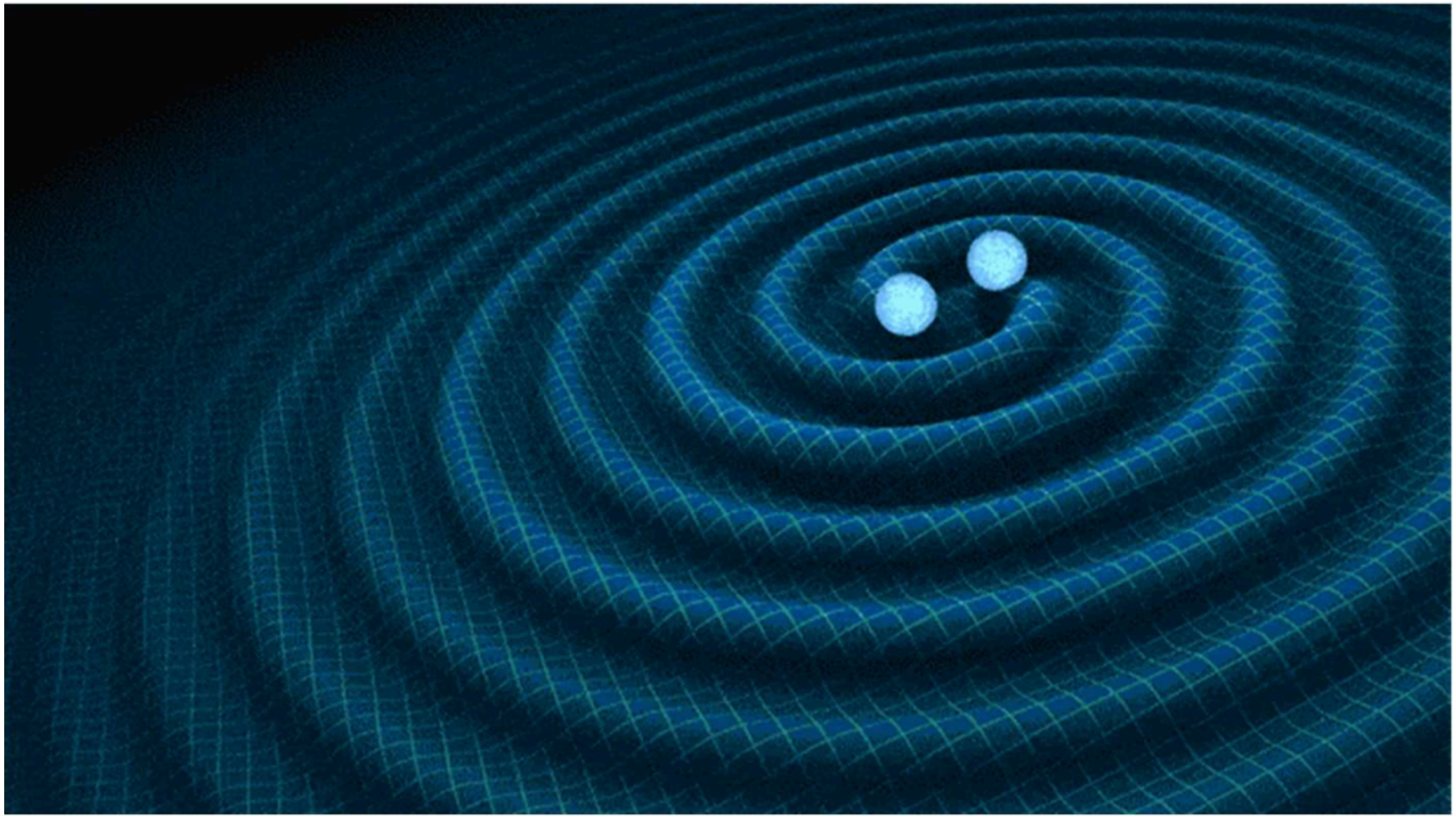


# Searching For Gravitational Waves

# Annoucnements

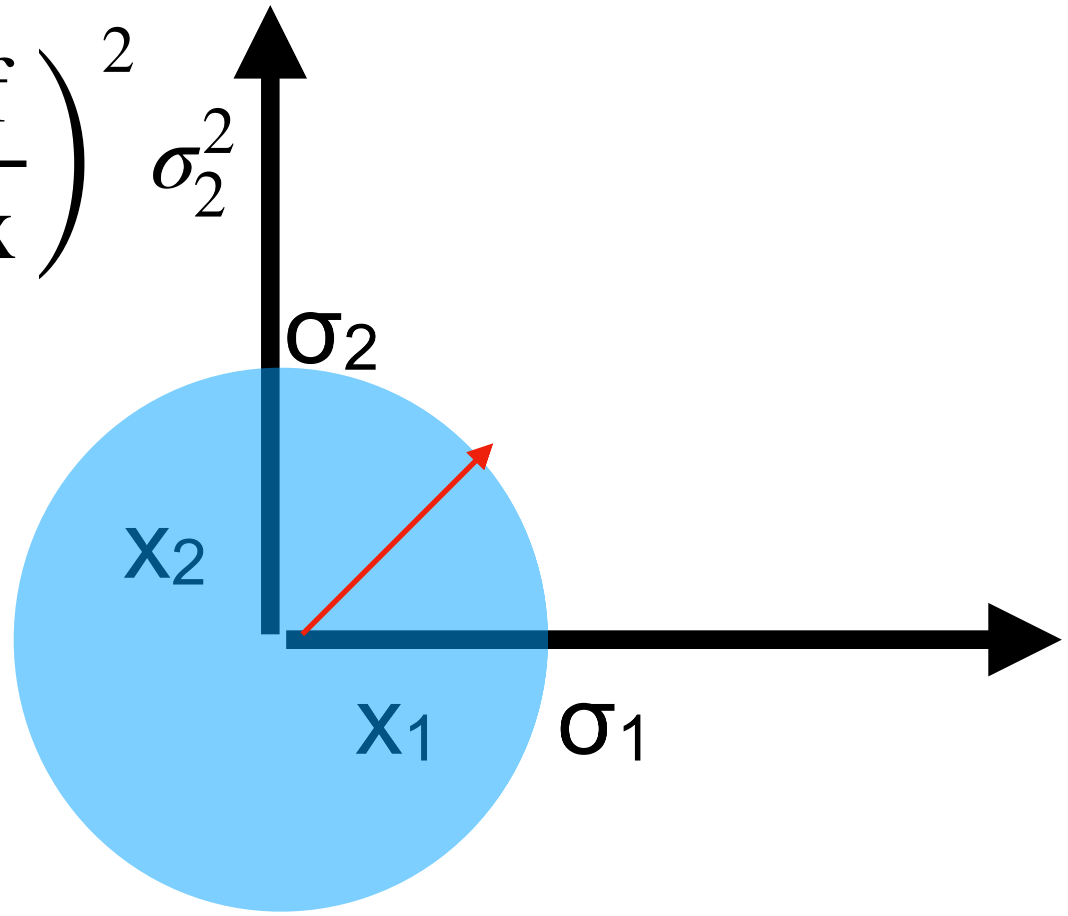
- You have a pset due this Friday at Midnight
  - Pset is available on canvas
  - Please send me/Sang Eon your comments and concerns

# Error Propagation

- An important element of variance is that it propagates
  - Take f(x)
  - What is the variance of f(x) given x
  - approx: $f(x+\Delta x) \approx f(x) + \Delta x \frac{df}{dx}$
  - Take (given $\Delta x \to \sigma$): $f(x+\sigma) \approx f(x) + \sigma \frac{df}{dx}$
  - $\mathrm{VAR[f(x)]} = \sigma^2 \left(\frac{\mathrm{df}}{\mathrm{dx}}\right)^2$

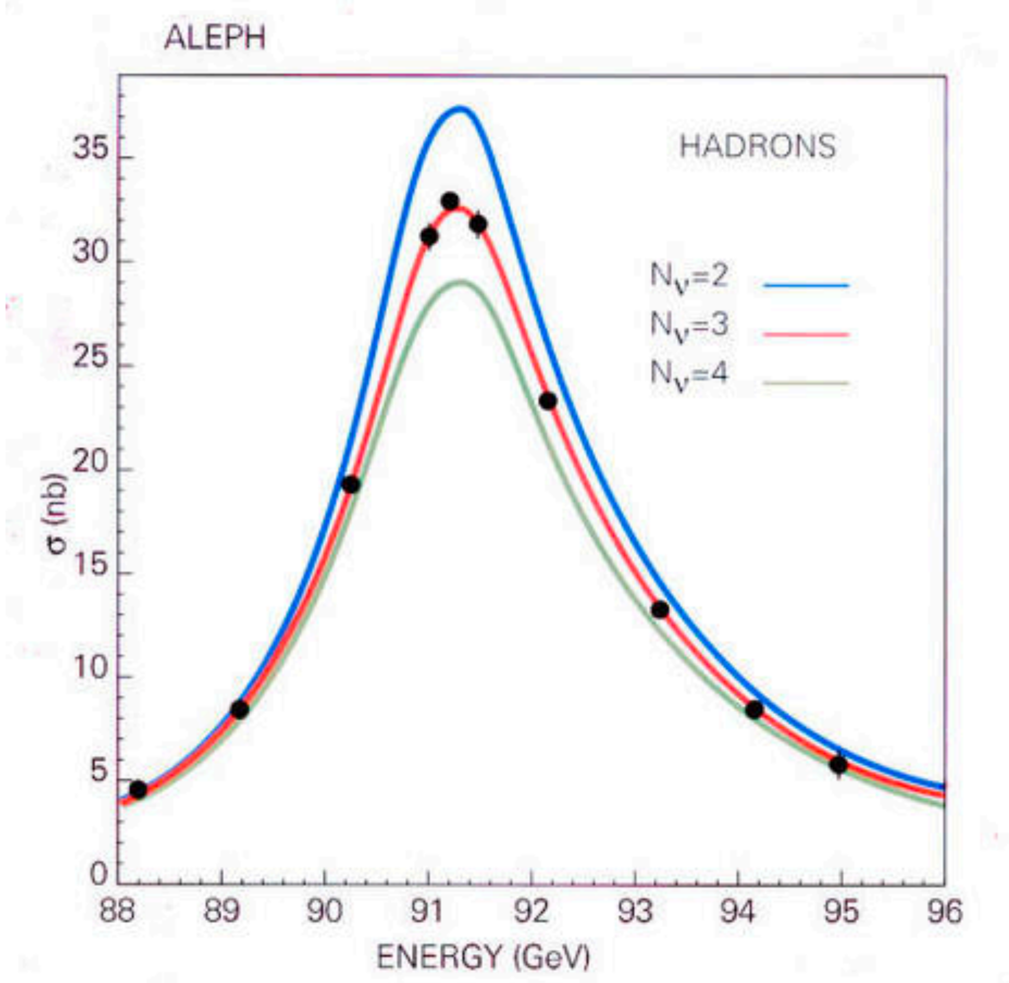
# What if you have two unc?

- Now consider another function:
- f(x) where $\sigma_1$ and $\sigma_2$ are independent uncertainties
- In this case we treat the total uncertainty as the sum
- Since they are independent, we can visualize them as a circle
- Take: $\mathrm{VAR}[f(x)] = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_1^2 + \left(\frac{\partial f}{\partial x}\right)^2 \sigma_2^2$



# Dealing with uncertainties

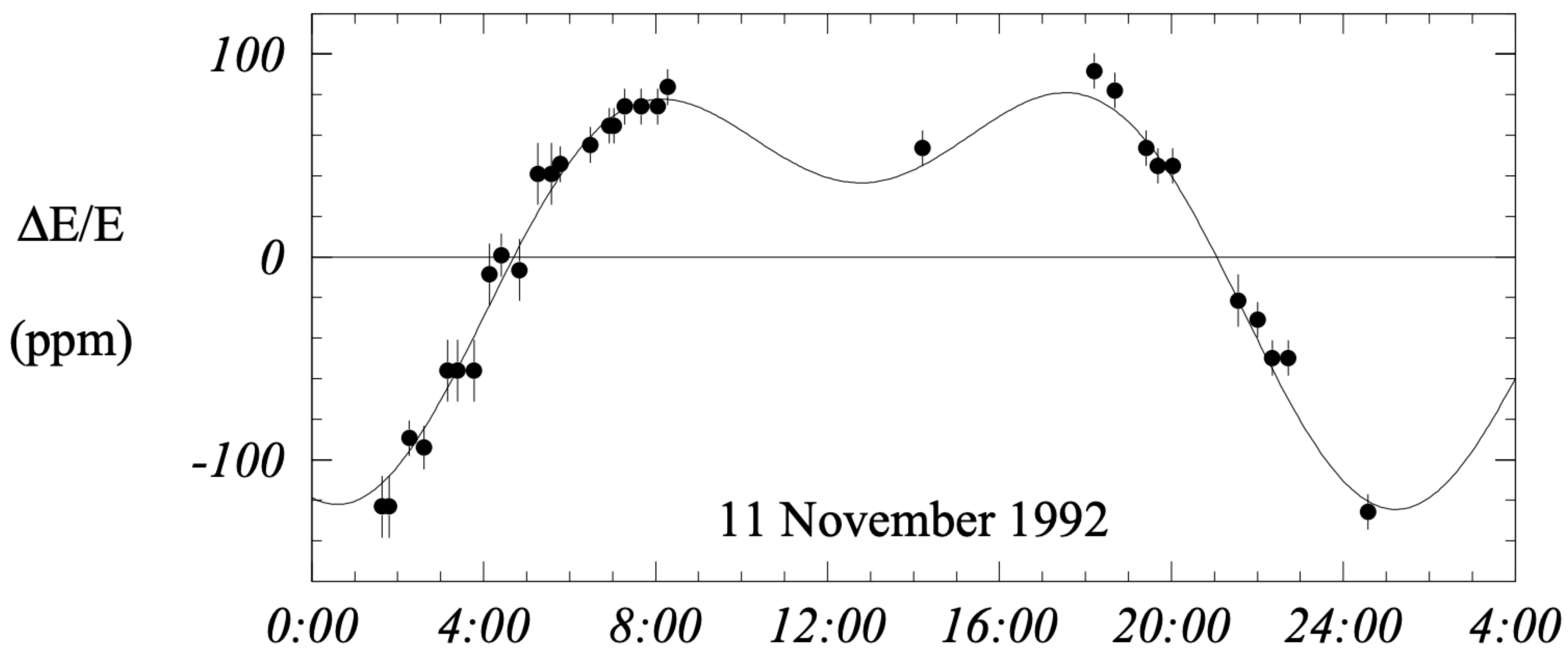
LEP: Large Elctron-Positron Collider




Collider was used to measure rates of collisions (and all their properties) Precisely

# Func with Unceratinties

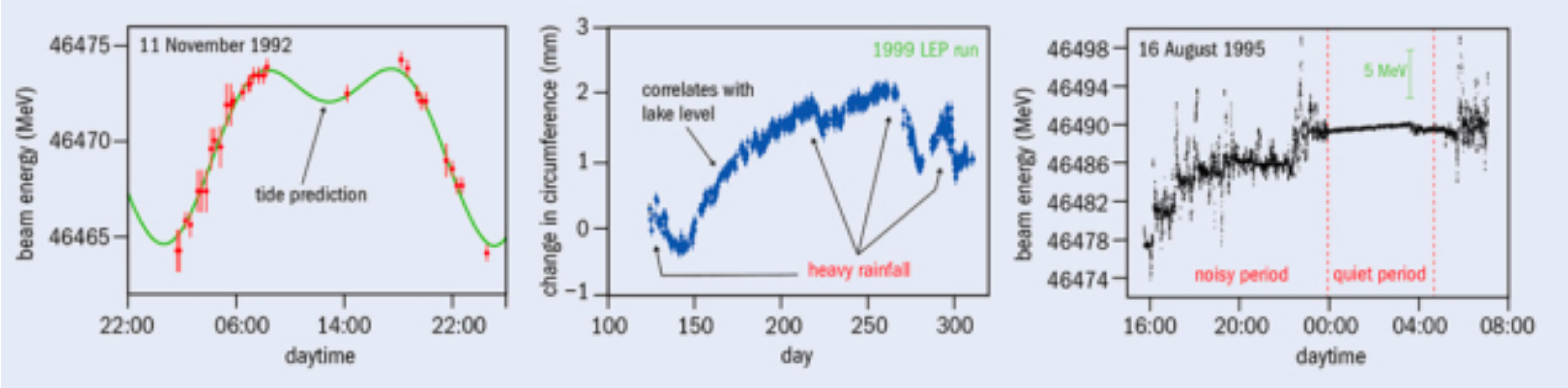
How precise is this?




Any guesses whats causing this?

Paper of this plot

# Lots of Uncertainties




- The collider would change parameters due to :
  - The water level in the nearby lake
  - The TGV train schedule
  - The orbits of the moon

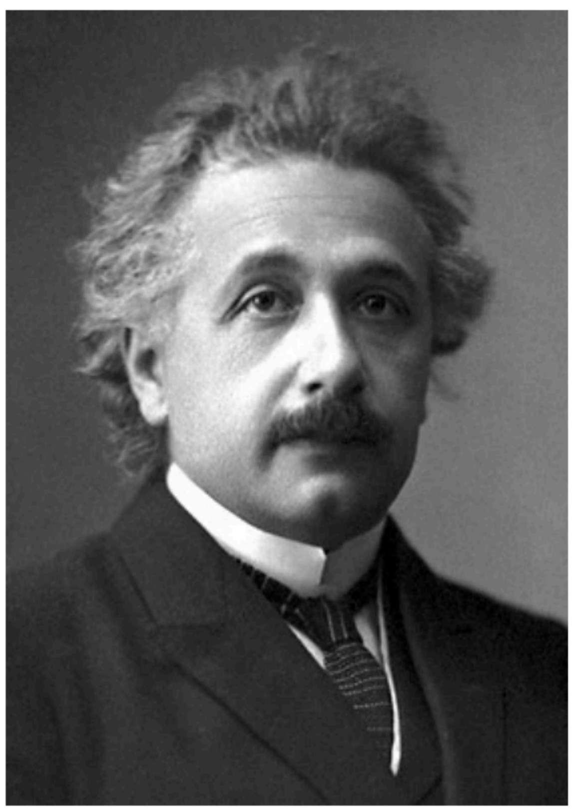
# Questions?

# Gravity

There is a limiting speed in Nature, the speed of light (1905)

geometry of spacetime

stress-energy tensor



$$\boxed{G_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta}}$$

Gravity: manifestation of spacetime curvature (1915)

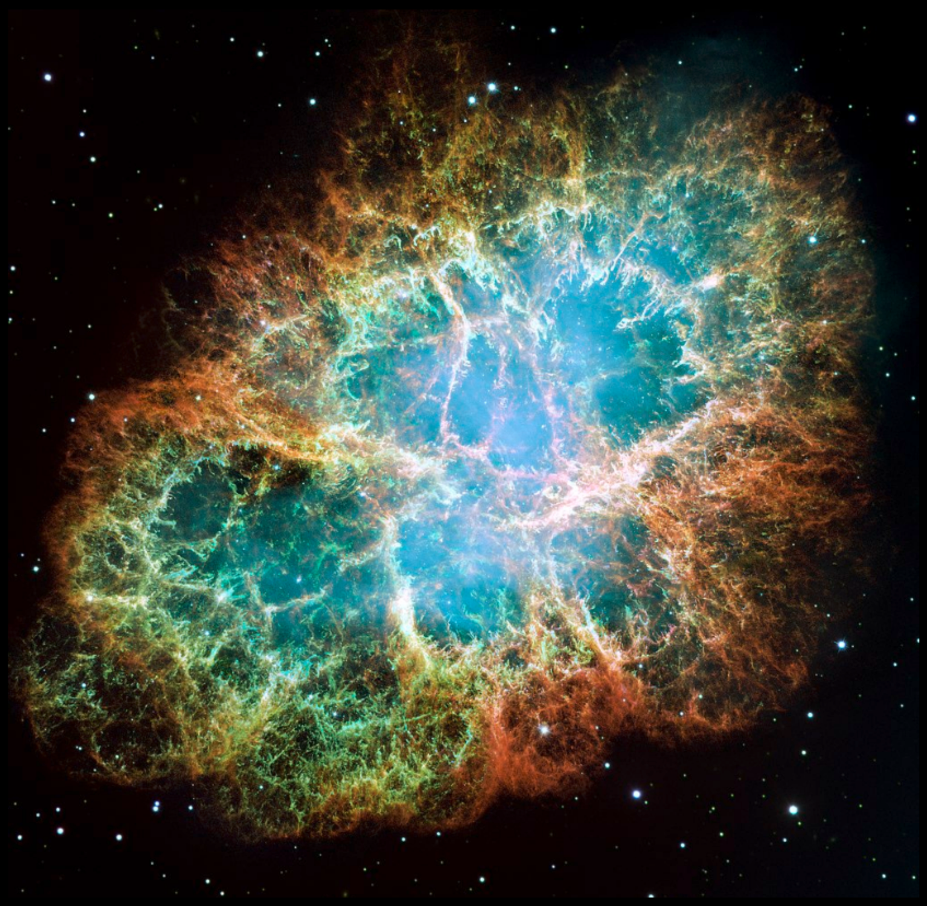
# Slides

- See this link for animations:
- https://www.dropbox.com/s/44ew35qlz89xono/PCH_Lecture2_LigoAnims_8S50.pptx?dl=0

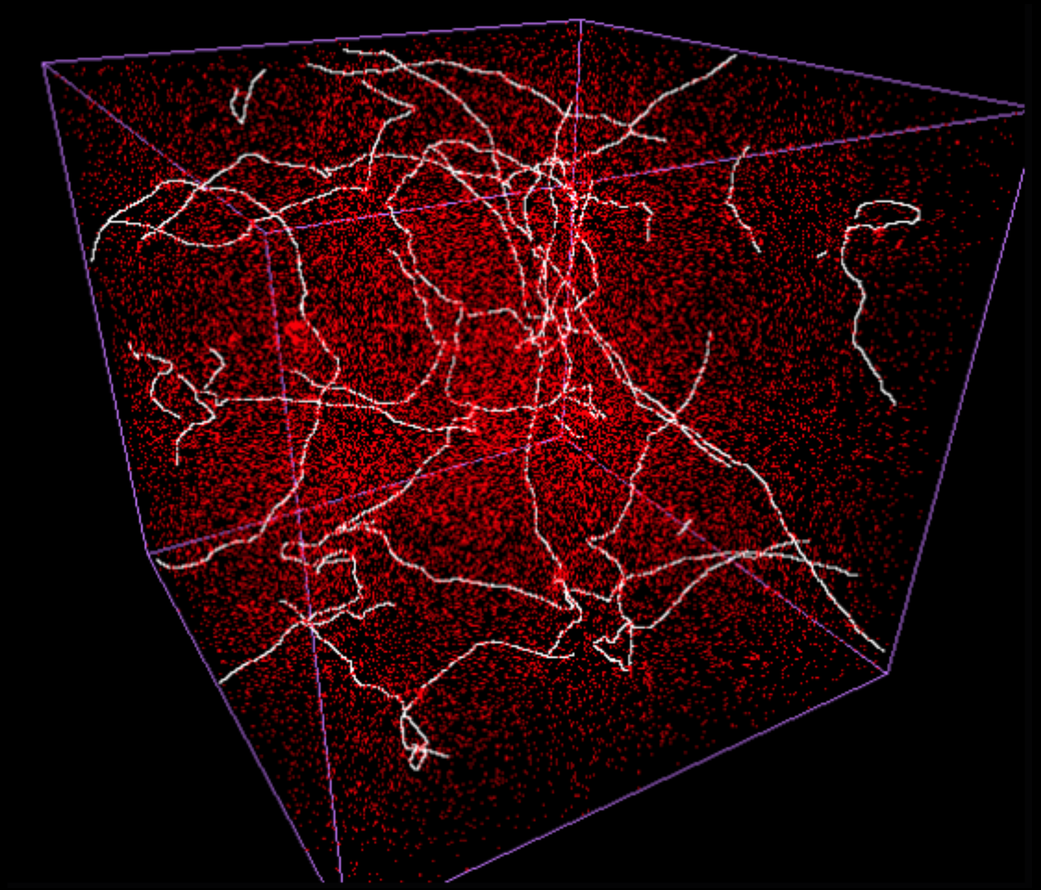
# Anomalous Gravitational Wave Sources

- There are other possible signal sources that can not be modelled and therefore cannot be detected using the match filtering pipeline
- We refer to them as anomalous and aim to develop a semi-supervised approach which would let us to discover such anomalous signals without explicit modelling

Core-collapse supernova (CCSN)

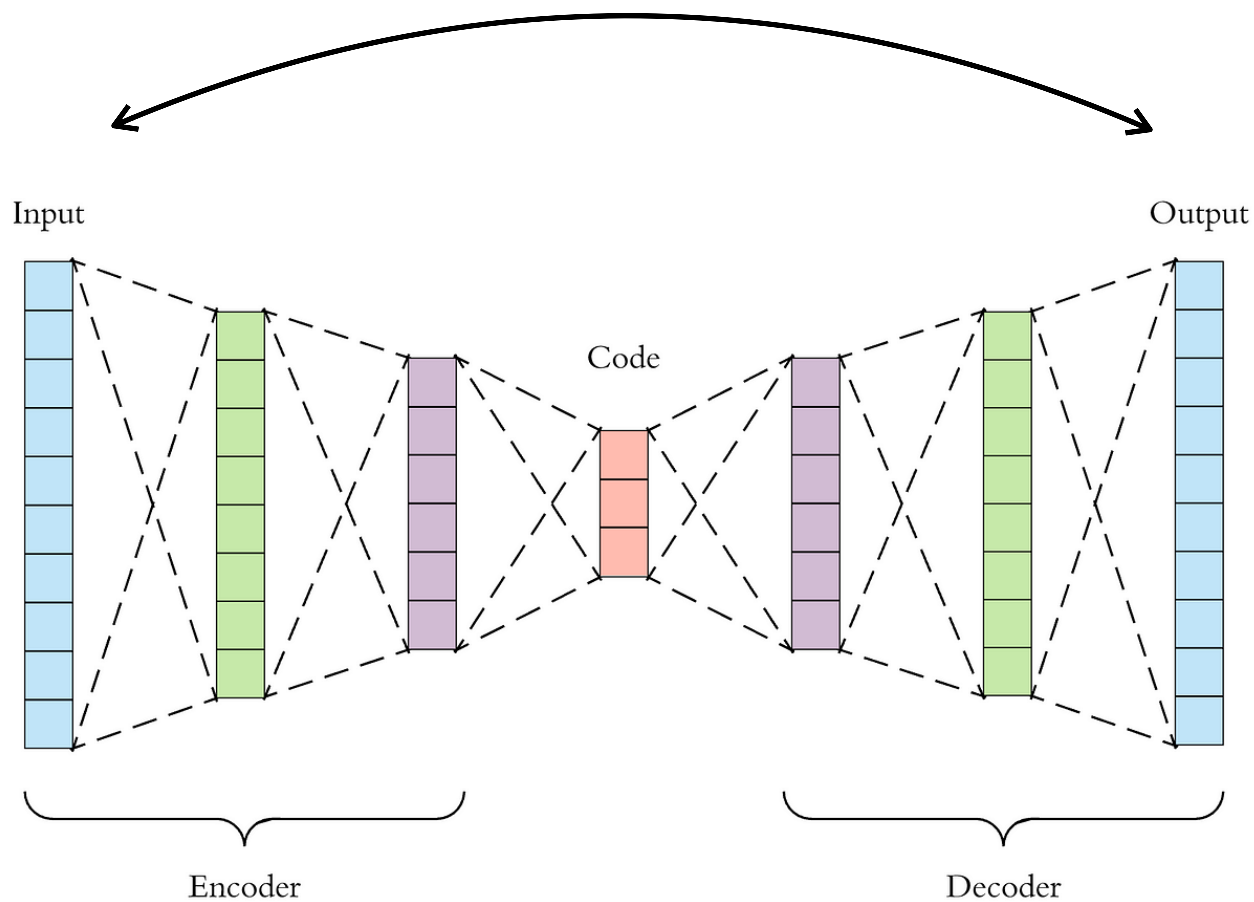


Cosmic strings



ARXIV2309.11537

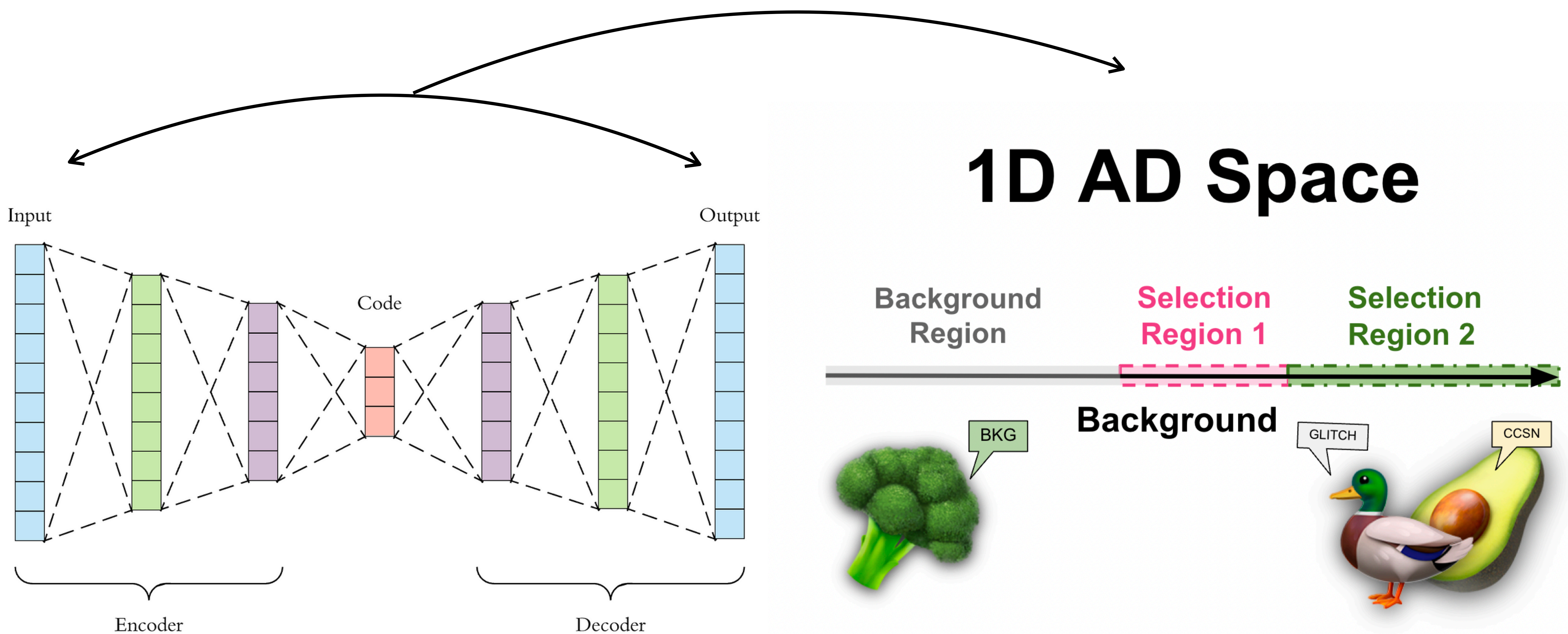
USE THE DISTANCE BETWEEN THE INPUT AND OUTPUT AS A METRIC FOR ANOMALY DETECTION



# Vanilla Anomaly Detection

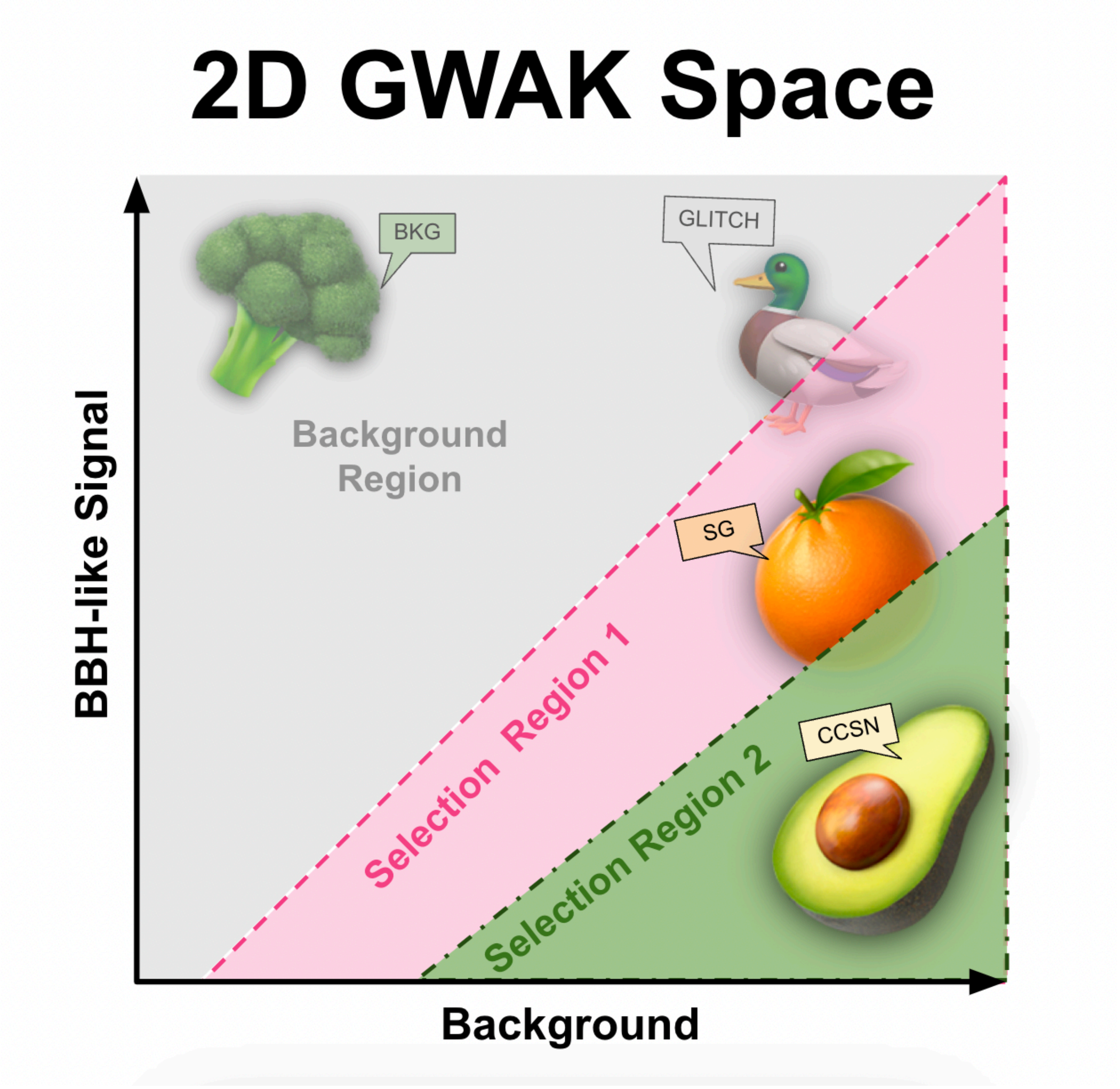
arxiv2309.11537

Use the distance between the input and output as a metric for anomaly detection



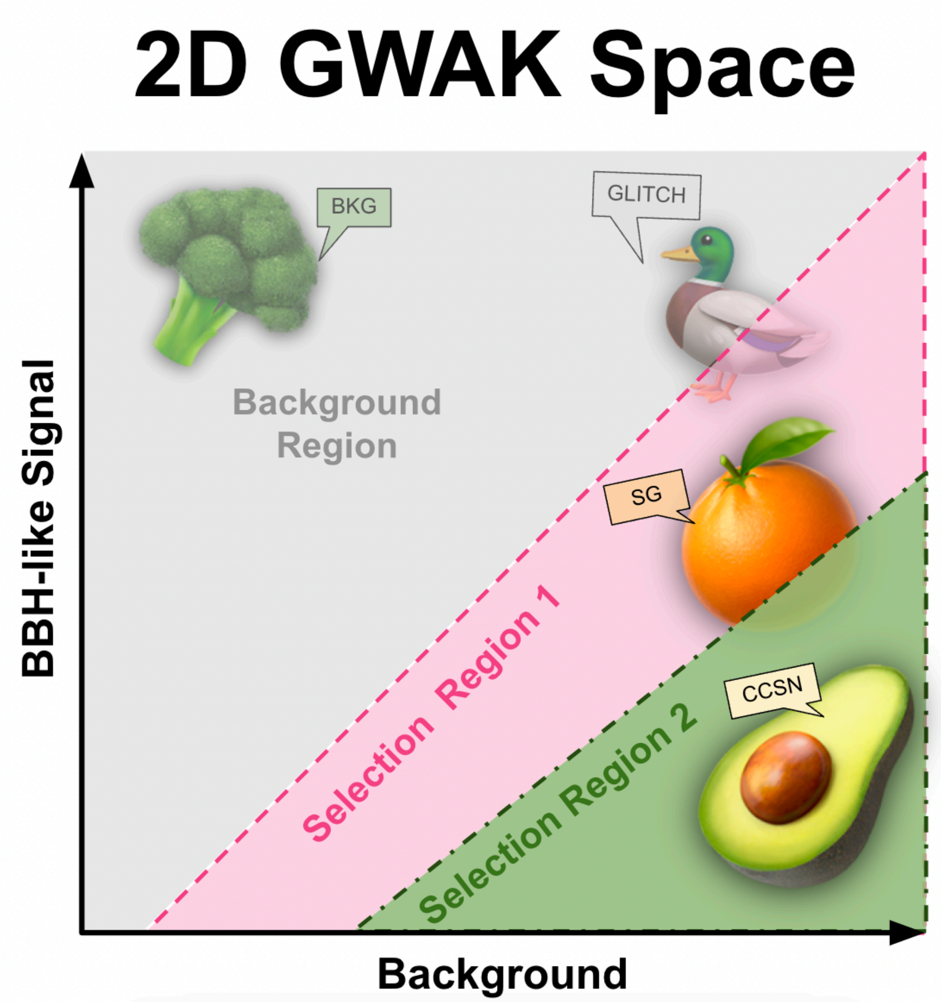
ARXIV2309.11537

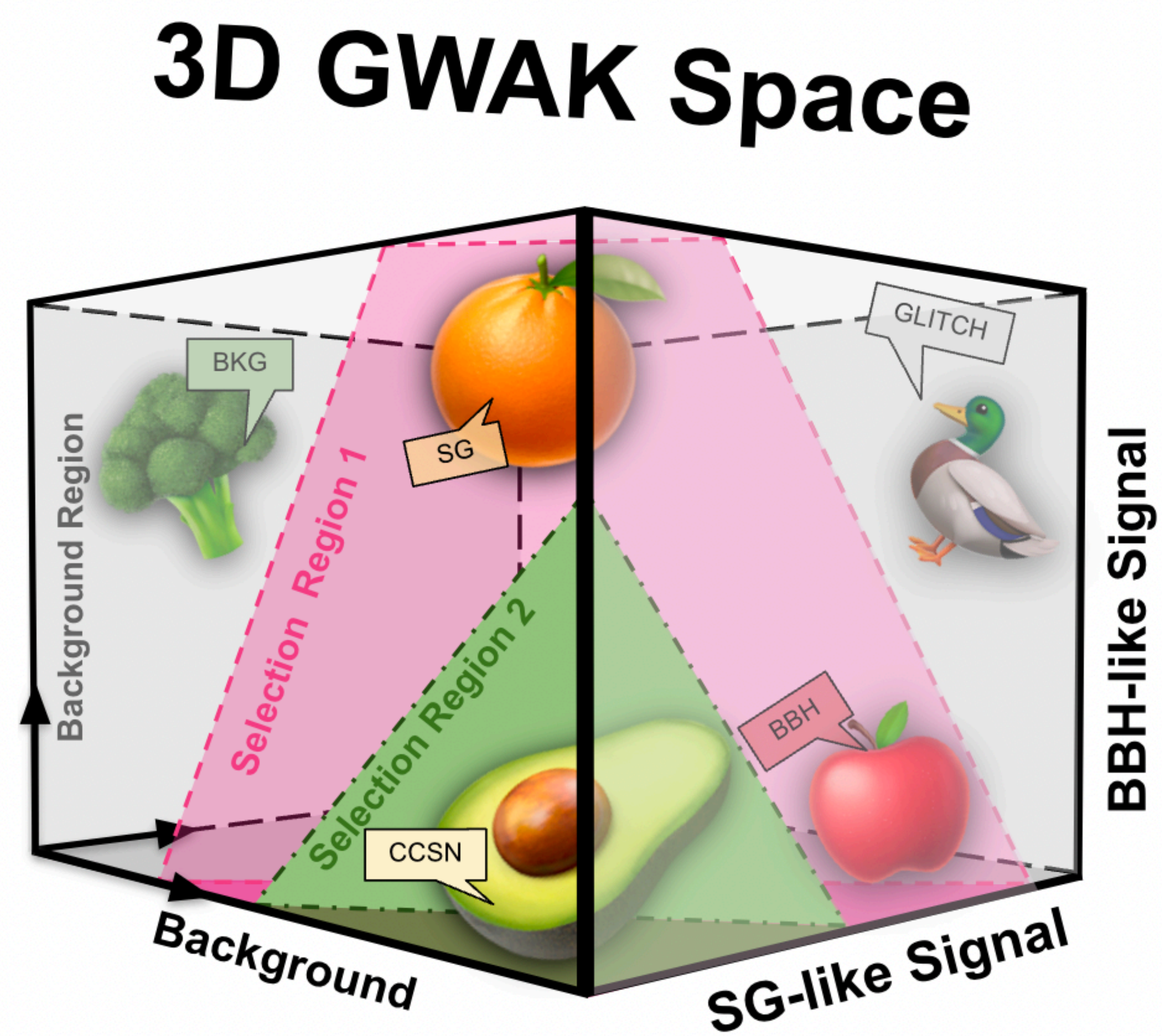
INCLUDING MORE AXES, BOTH SIGNAL AND BACKGROUND, ALLOWS TO MORE EFFICIENTLY SELECT A SIGNAL-LIKE ANOMALIES



ARXIV2309.11537

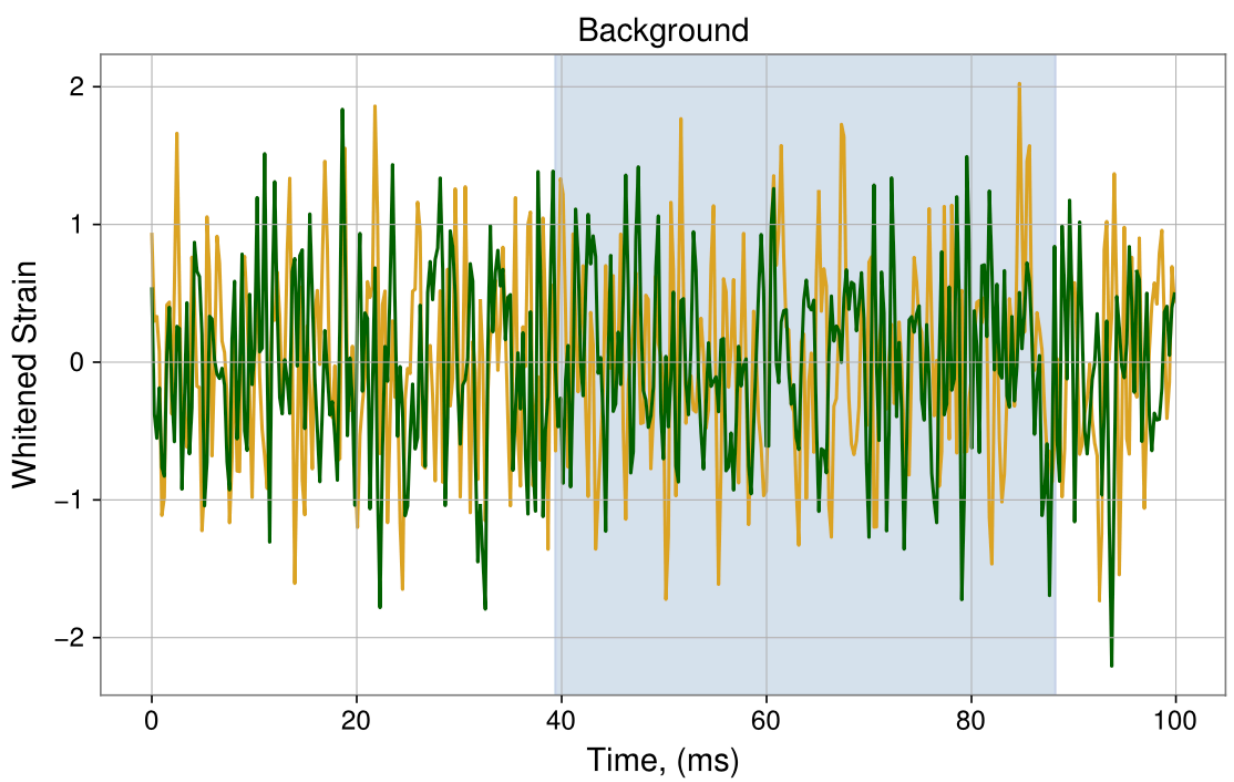
INCLUDING MORE AXES, BOTH SIGNAL AND BACKGROUND, ALLOWS TO MORE EFFICIENTLY SELECT A SIGNAL-LIKE ANOMALIES

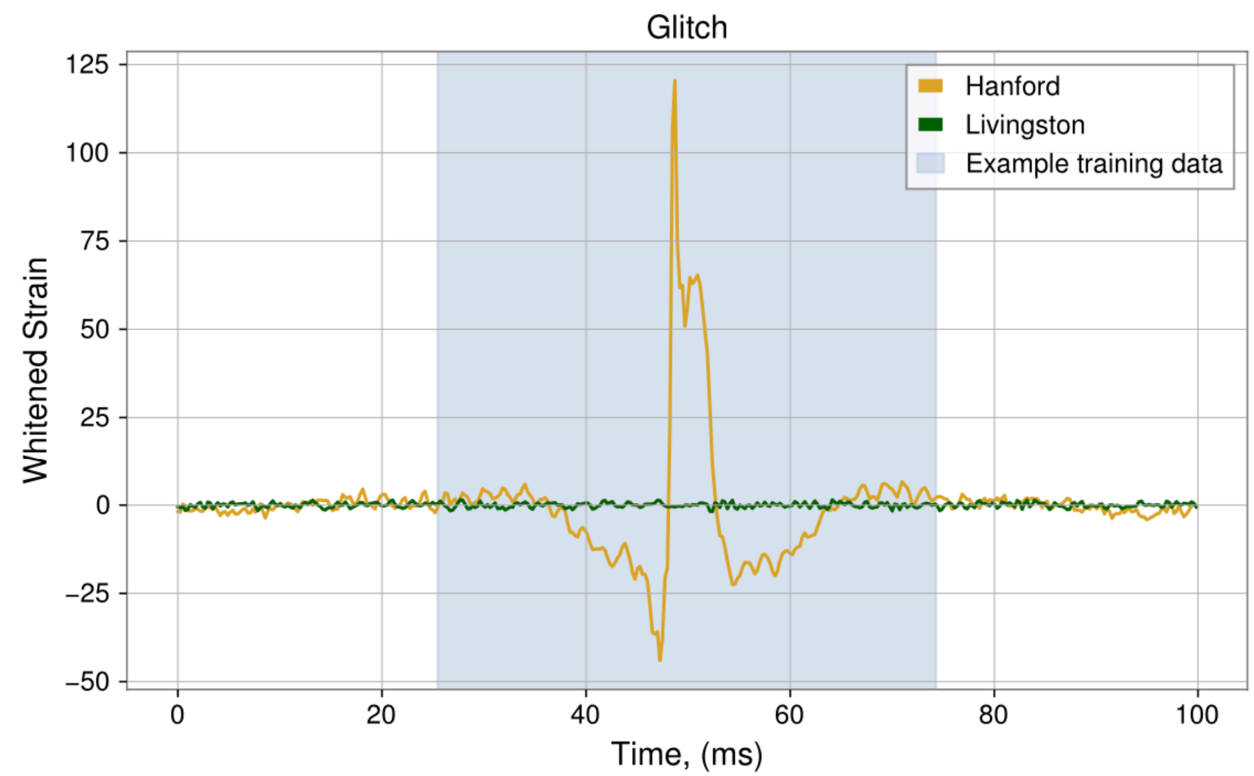


EXAMPLE OF GWAK CLASSES: GLITCH AND BACKGROUND STRAINS

THE LIGHT BLUE SHADING HIGHLIGHTS AN EXAMPLE REGION THAT IS PASSED AS INPUT TO THE AUTOENCODERS FOR TRAINING

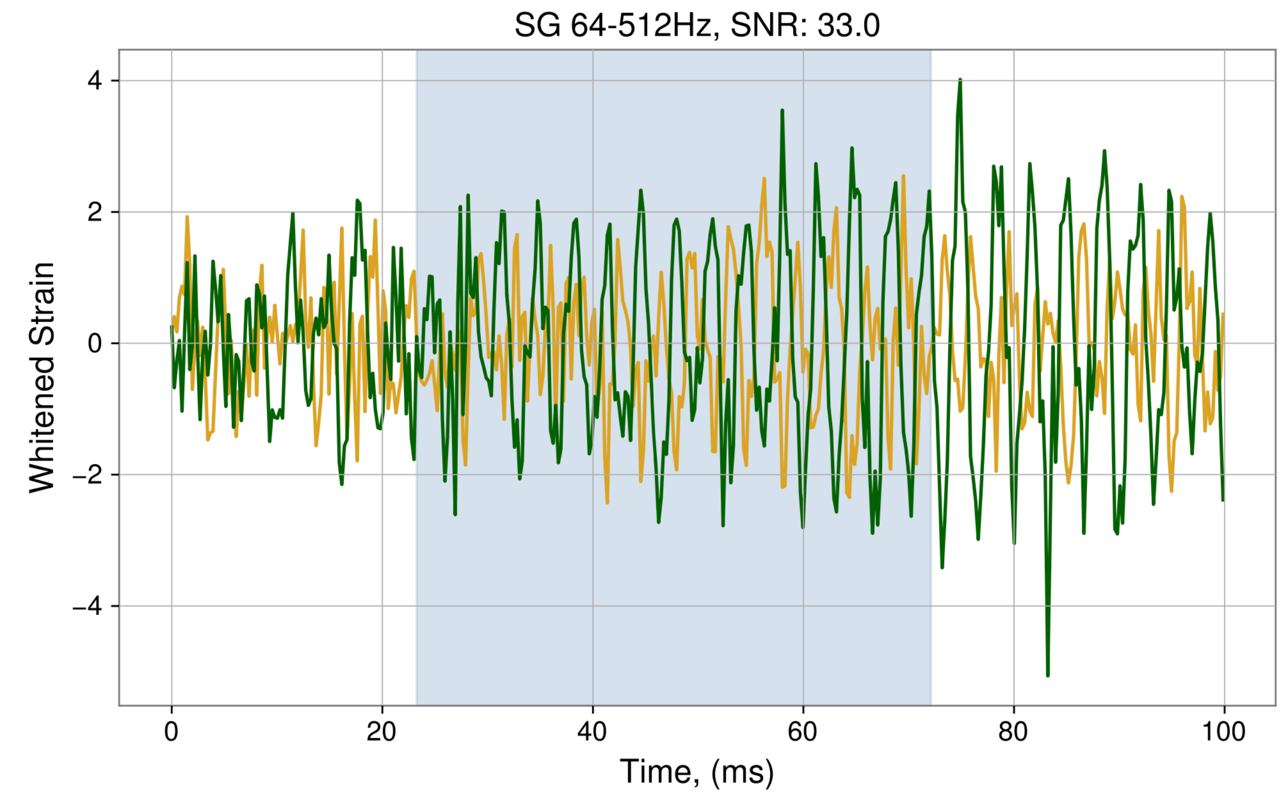


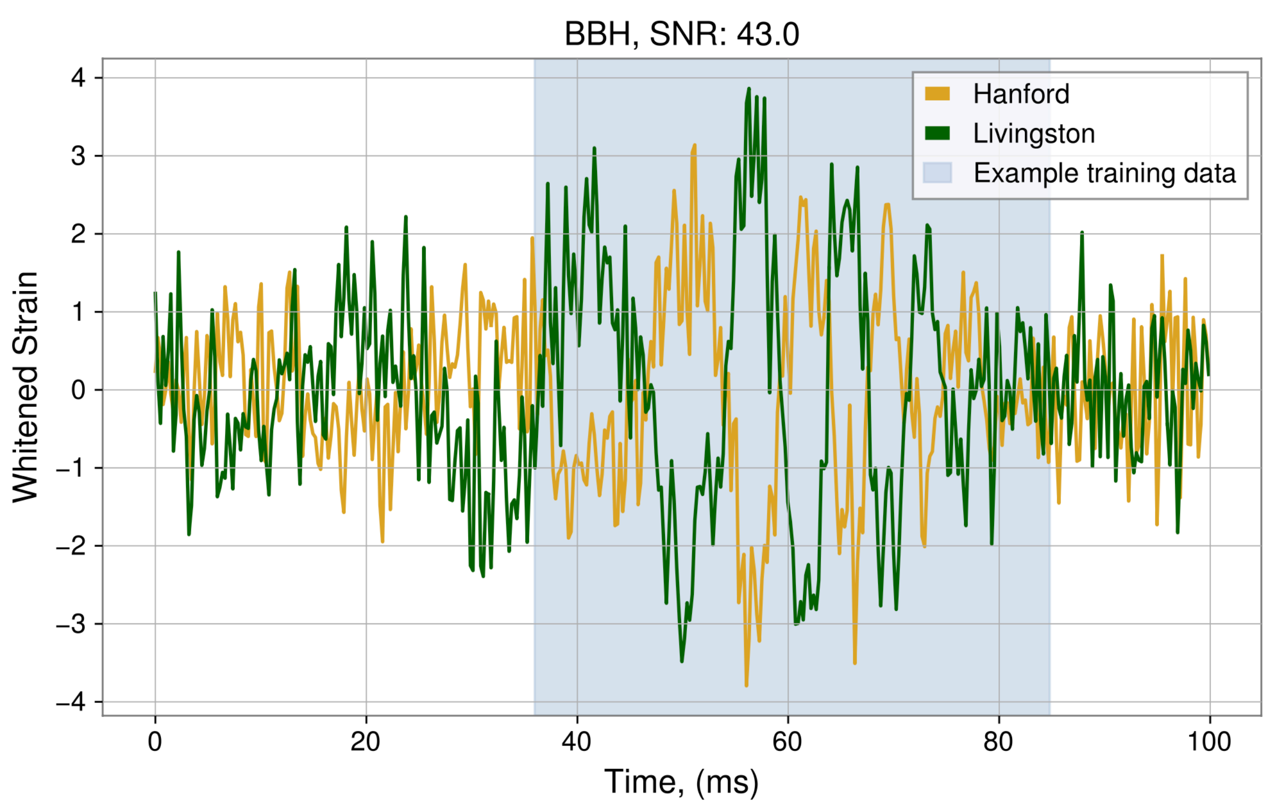



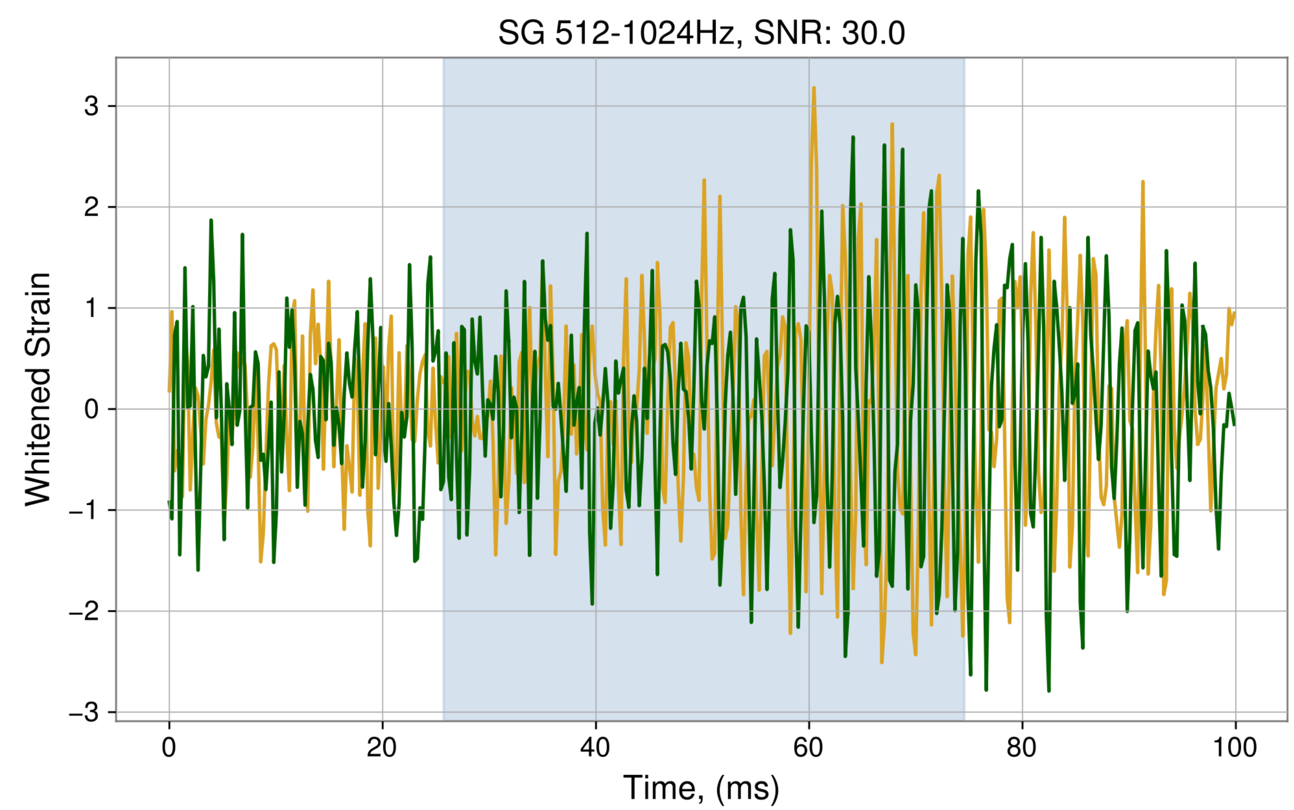
ARXIV2309.11537

EXAMPLE OF SIGNAL-LIKE CLASSES: BBH AND SINE-GAUSSIAN STRAINS FROM LIVINGSTON AND HANFORD

THE LIGHT BLUE SHADING HIGHLIGHTS AN EXAMPLE REGION THAT IS PASSED AS INPUT TO THE AUTOENCODERS FOR TRAINING



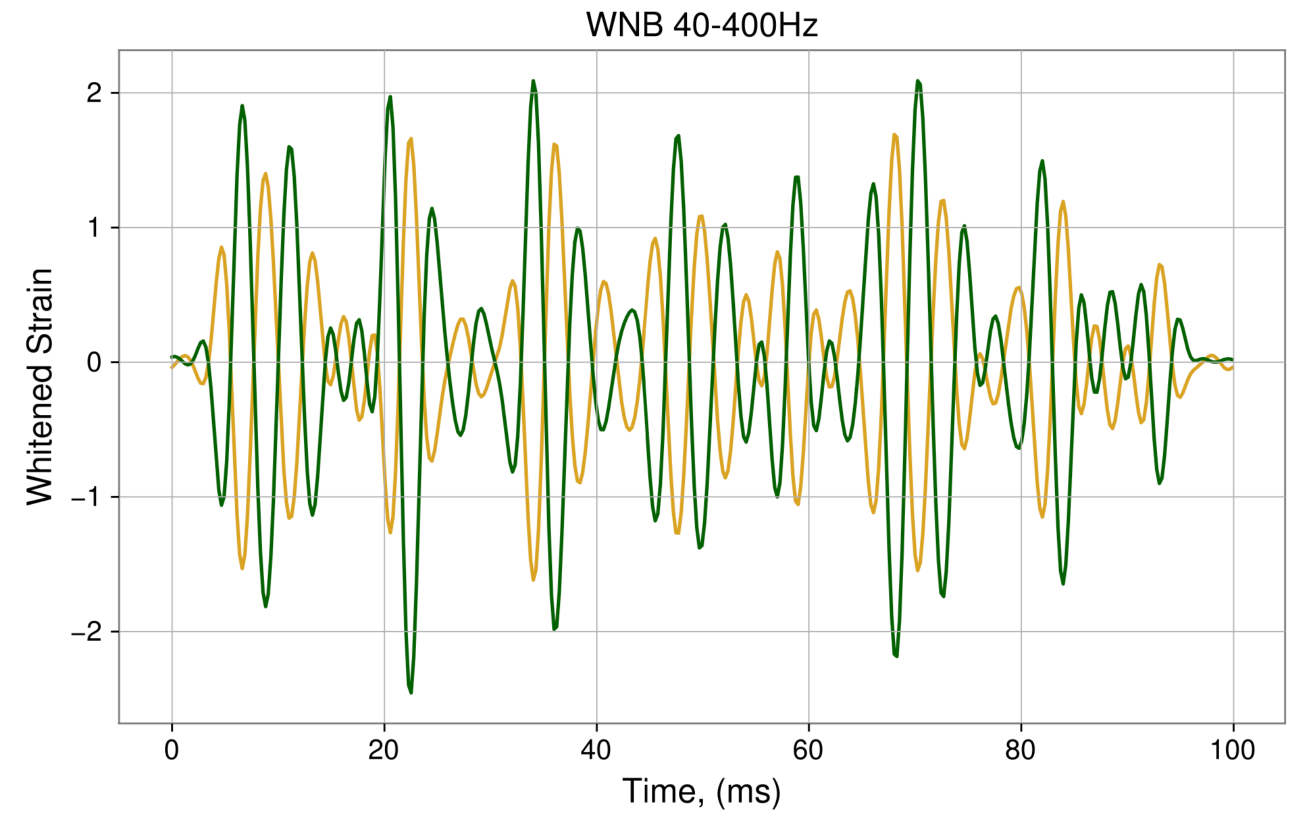



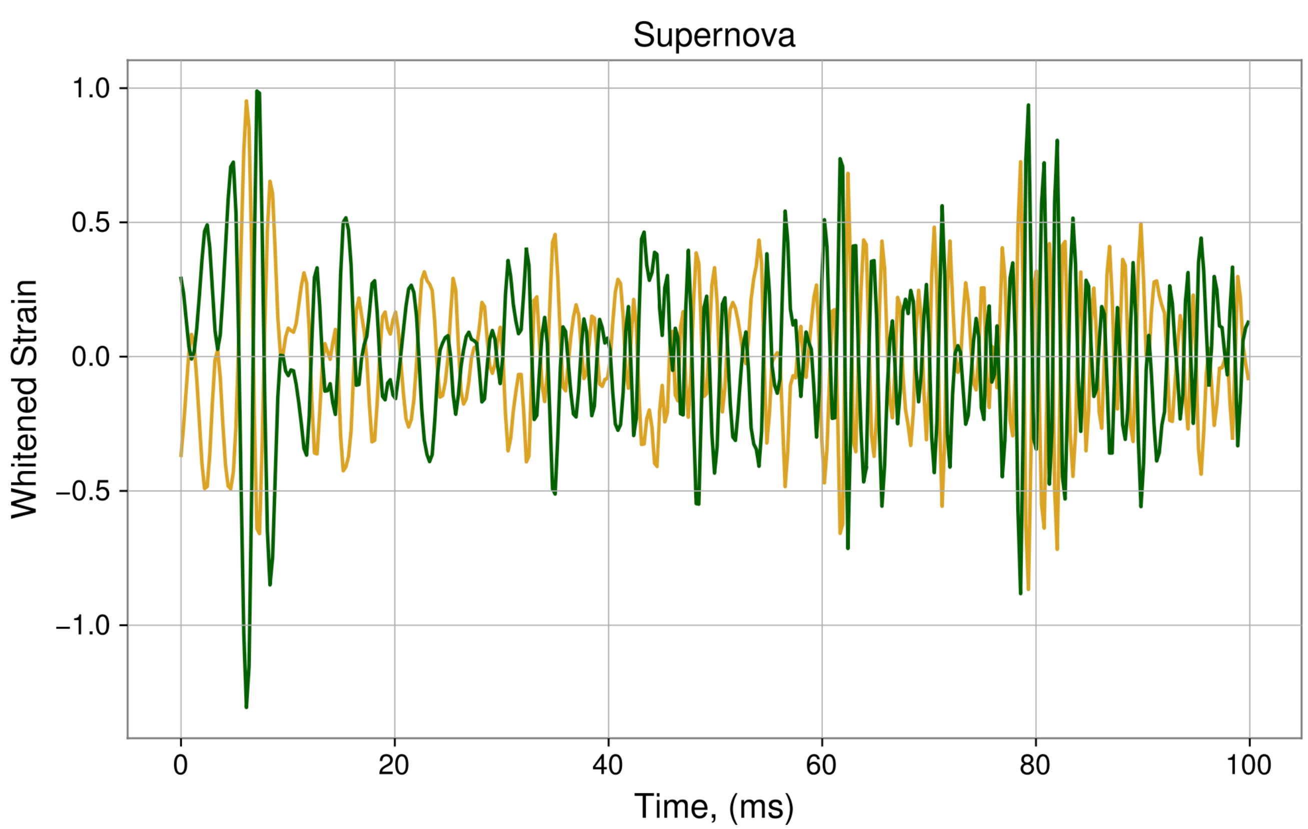



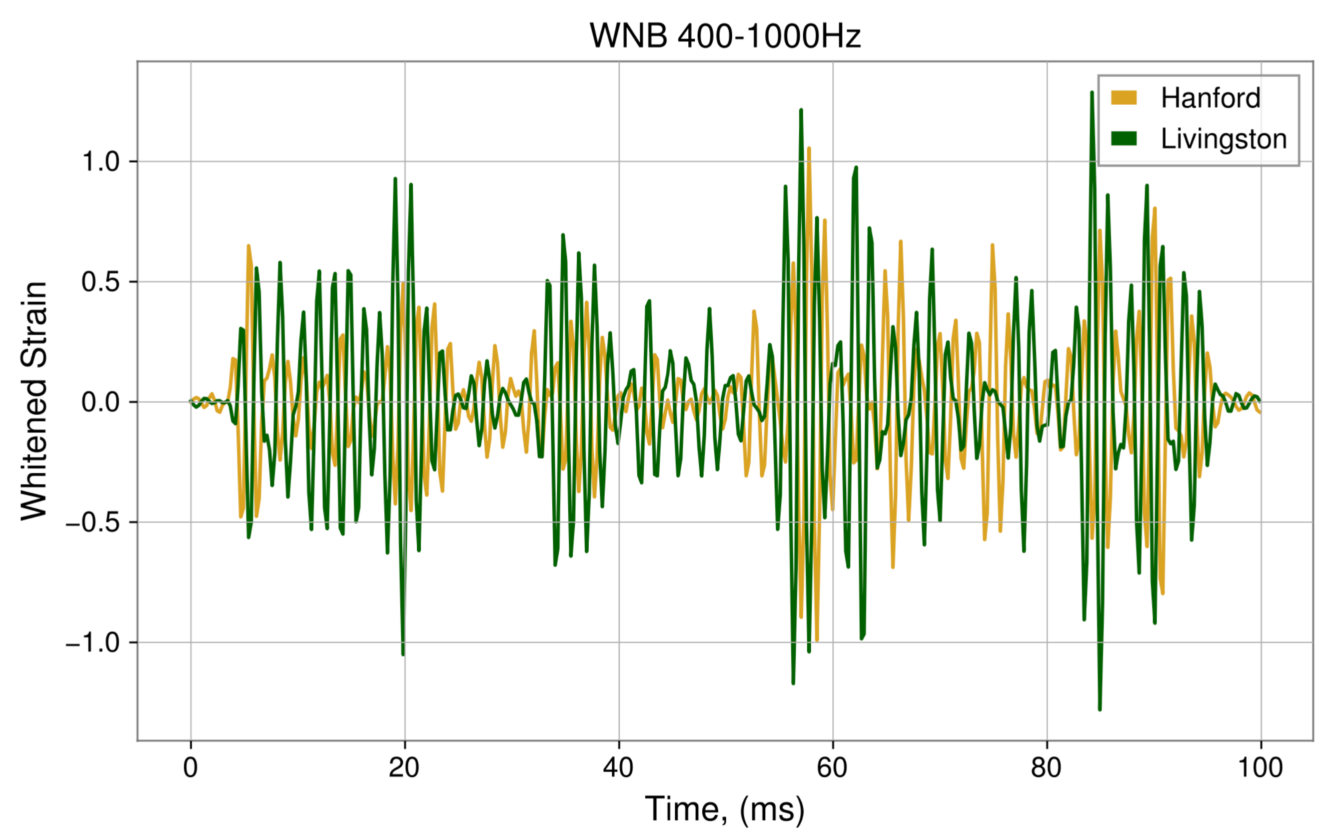
ARXIV2309.11537

EXAMPLE OF SIGNAL-LIKE CLASSES: SUPERNOVA AND WHITE NOISE BURST STRAINS FROM LIVINGSTON AND HANFORD

THOSE ANOMALIES ARE NOT USED TO CREATE THE GWAK



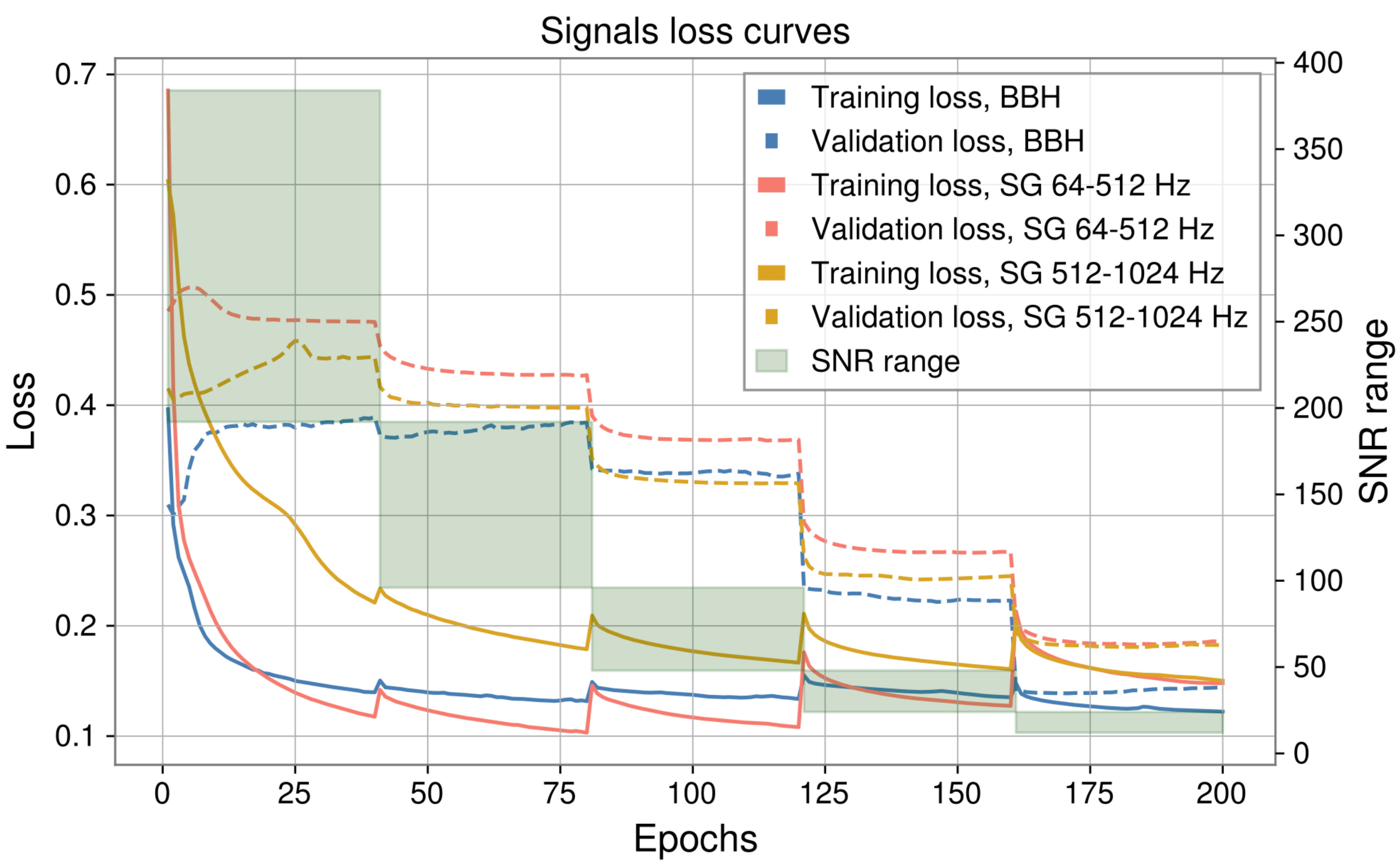
arXiv2309.11537

Autoencoder training and validation losses for signal classes, using curriculum learning to progressively reduce validation loss

The validation loss for each training SNR range is computed on the validation data from the last SNR step

In solid/dashed colours are the training/validation losses for BBH, SG 64-512 Hz and SG 512-1024 Hz

A light green shaded region depicts the SNR range for each step of training, spanning the range of injected SNR corresponding to each curriculum.

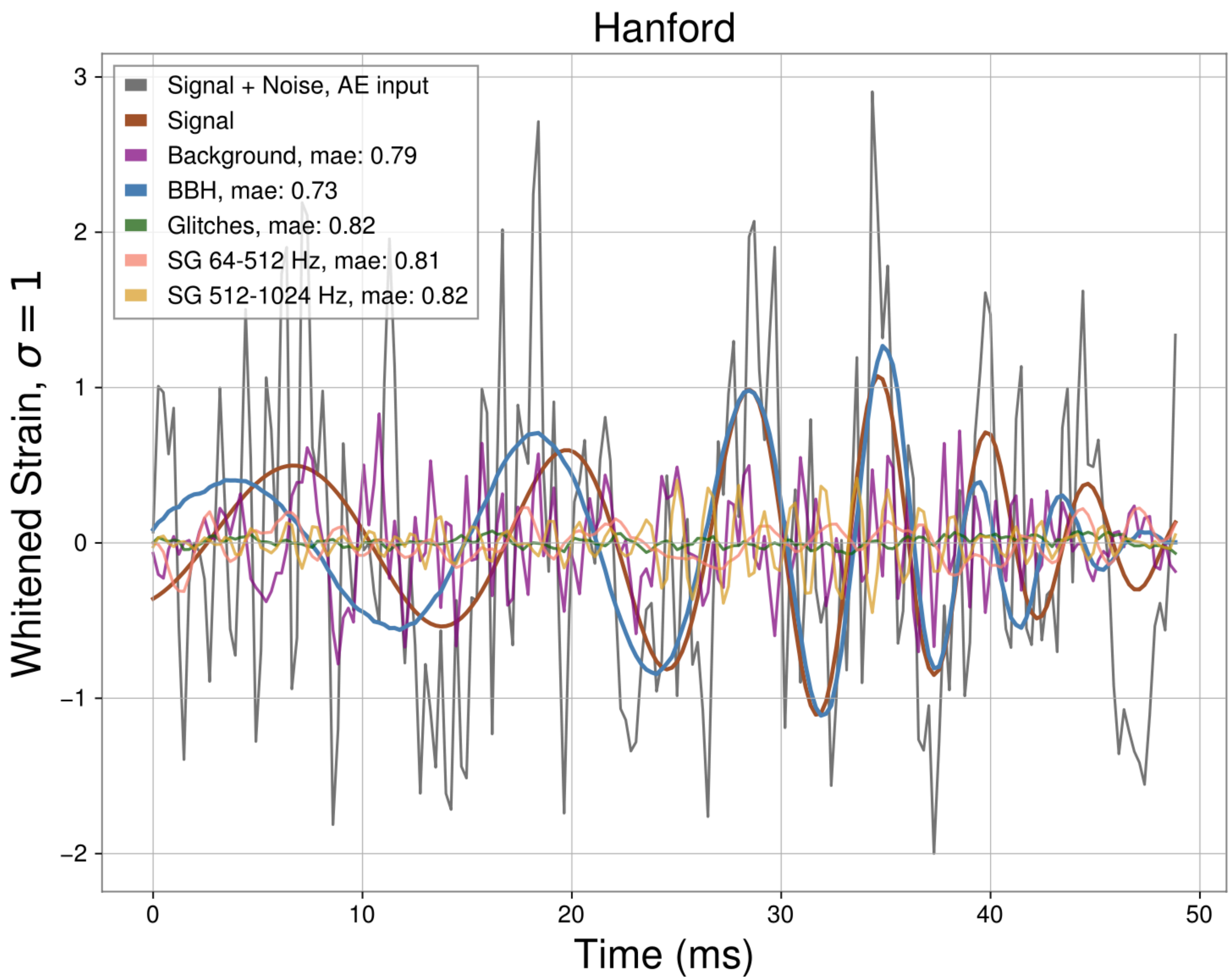


ARXIV2309.11537

EXAMPLE OF RECREATION ON INJECTED BBH SIGNAL, WITH THE NOISE-LESS TEMPLATE SHOWN AS WELL

THE RECREATION OF THE BBH AUTOENCODER FOLLOWS CLOSELY THE ORIGINAL SIGNAL INJECTION

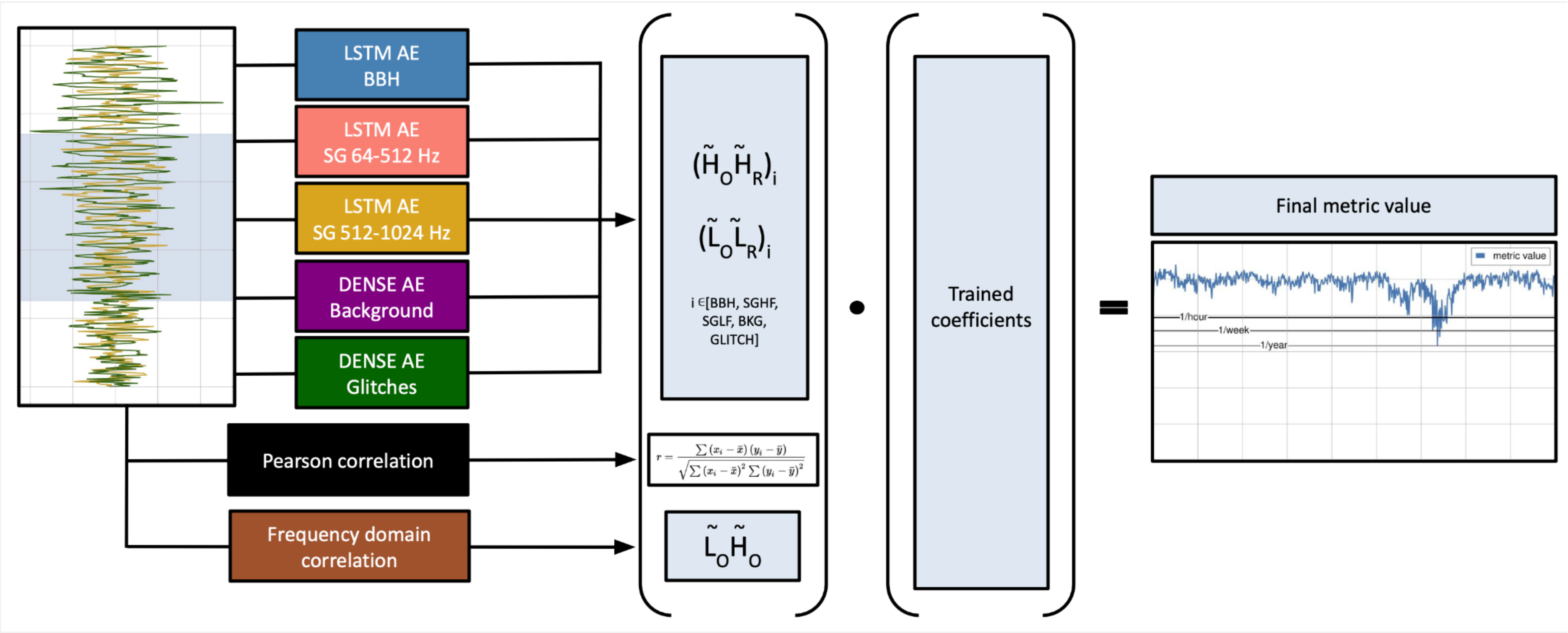
WHILE BACKGROUND, GLITCHES, SG 64-512 HZ AND SG 512-1024 HZ FAIL TO RECONSTRUCT THE INJECTED BBH SIGNAL



ARXIV2309.11537

EXAMPLE OF METRIC CALCULATION ON A HYPOTHETICAL EVENT

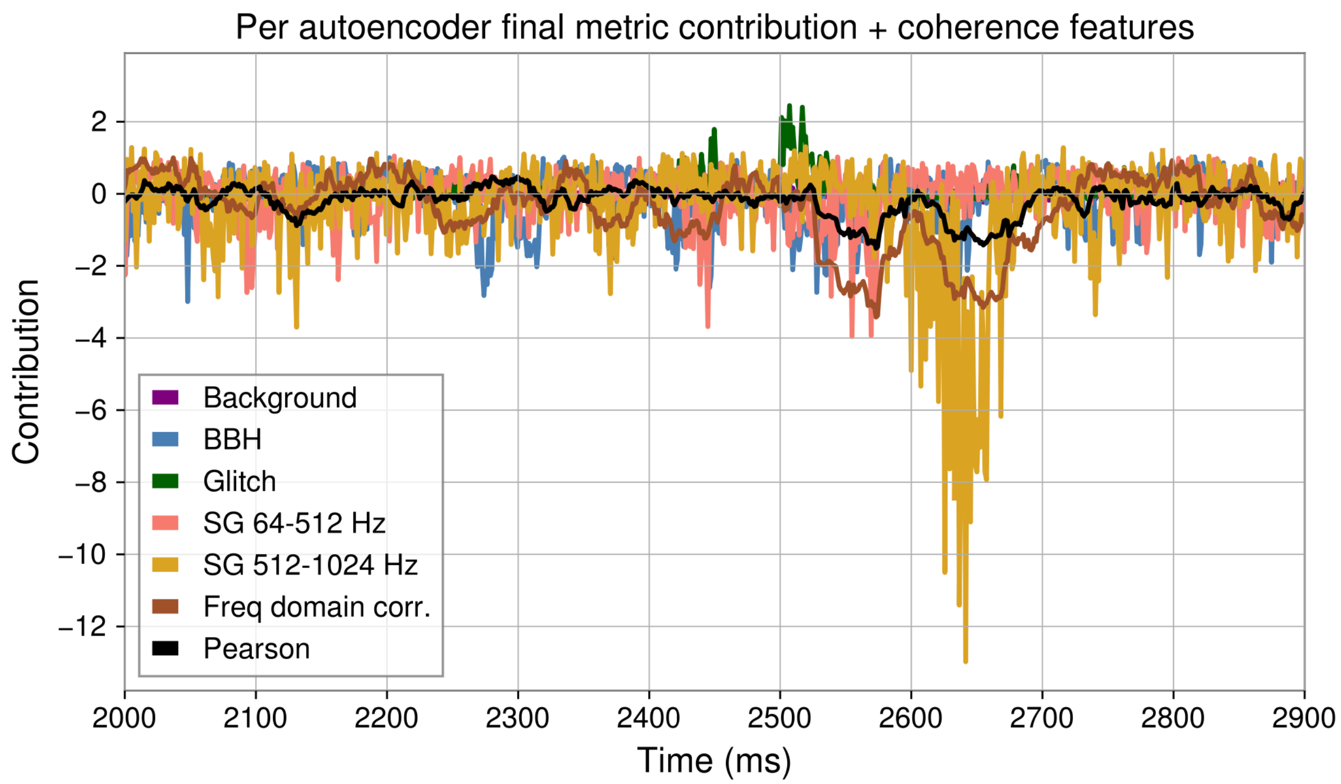
1. THE EVENT IS RECONSTRUCTED WITH EACH OF THE 5 PRE-TRAINED AUTOENCODERS, 2 FEATURES PER AE
2. PEARSON AND FREQUENCY DOMAIN CORRELATION ARE COMPUTED ON THE GIVEN INPUT
3. EACH OF THE VALUES IS MULTIPLIED WITH A CORRESPONDING COEFFICIENT WHICH ARISES FROM THE PRE-TRAINED LINEAR METRIC
4. THE SUM OF ALL THE FEATURES MULTIPLIED BY THEIR COEFFICIENTS IS REFERRED TO AS THE FINAL METRIC AND IS USED TO MAKE A DECISION ON IF THE EVENT IS SIGNAL-LIKE

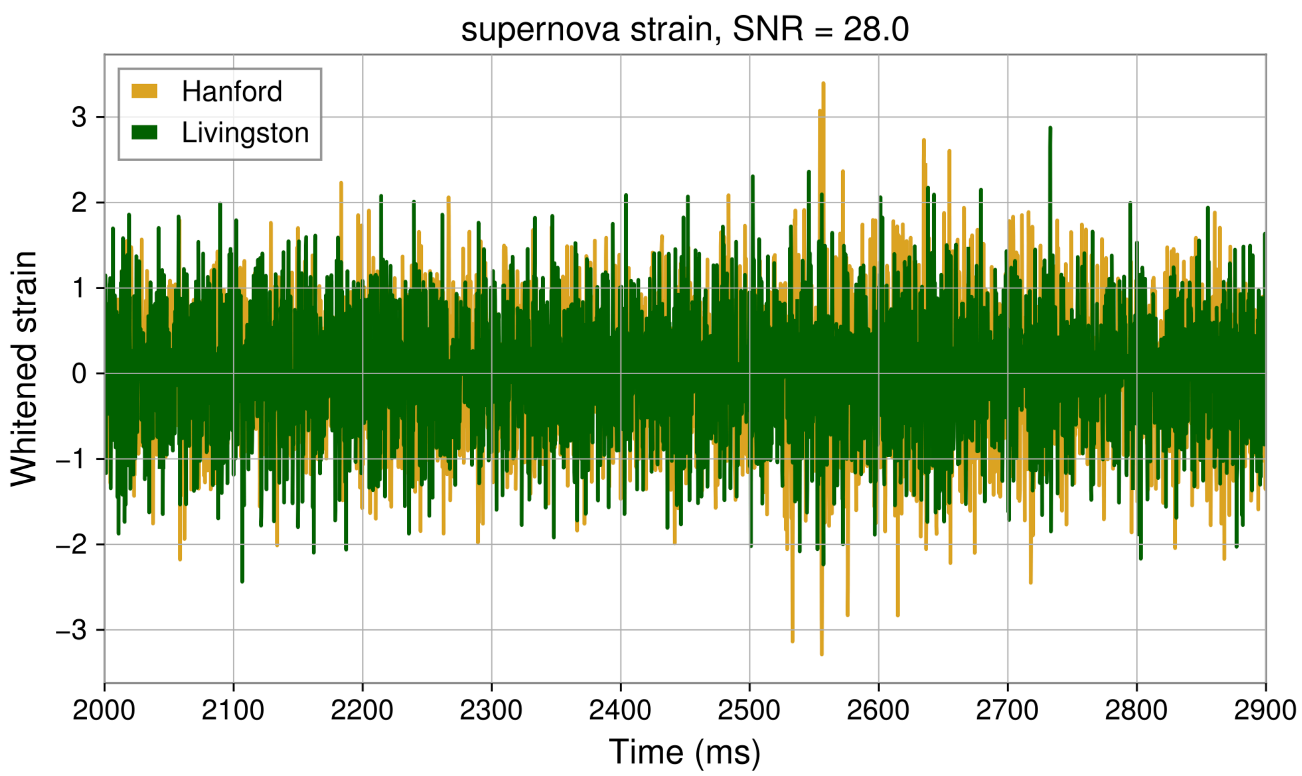


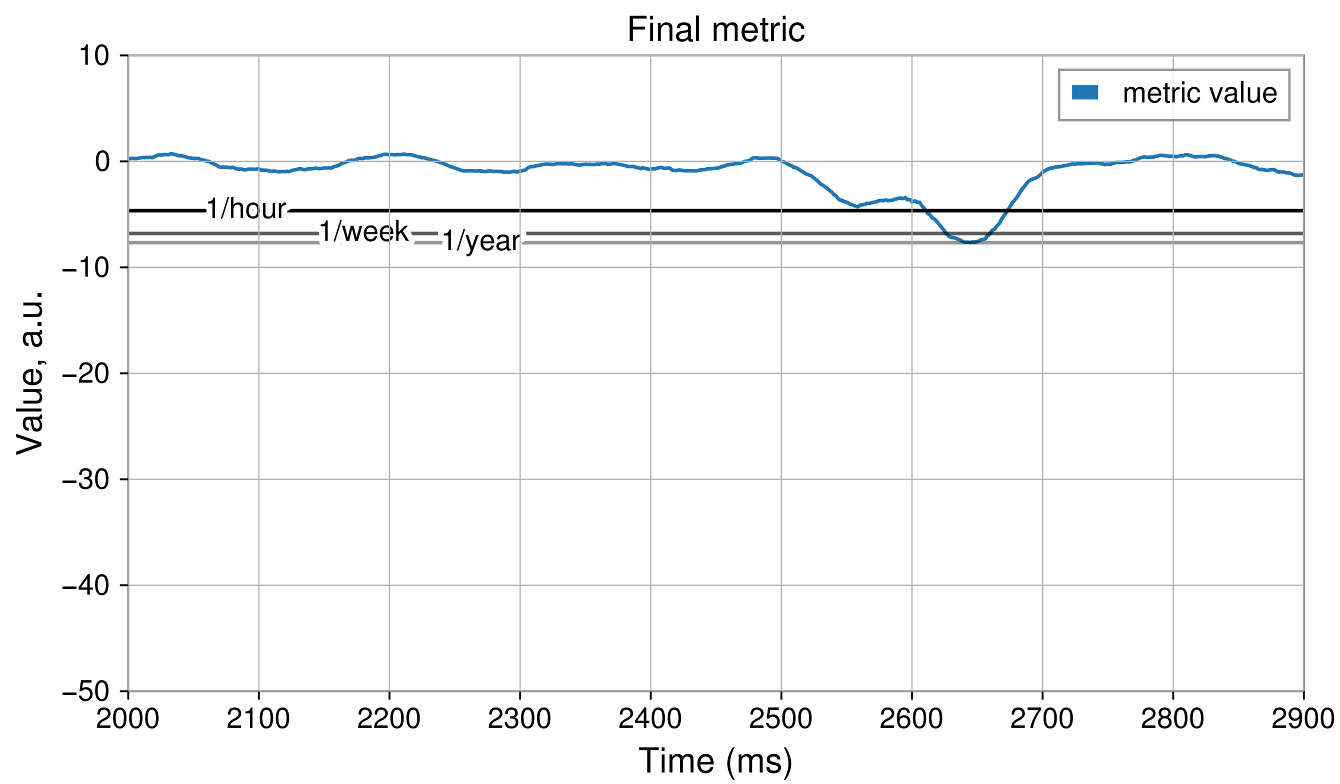
ARXIV2309.11537

STRAIN, GWAK METRIC RESPONSE AND FINAL METRIC RESPONSE FOR SUPERNOVA SIMULATED SIGNAL

THE EVALUATION OF GWAK AXES AND PEARSON CORRELATION WITH TIME AND ON THE TOP RIGHT TOTAL METRIC VALUE AND FAR ARE SHOWN AS AN EXAMPLE OF THE ALGORITHM'S "REACTION" TO AN UNSEEN SIGNAL

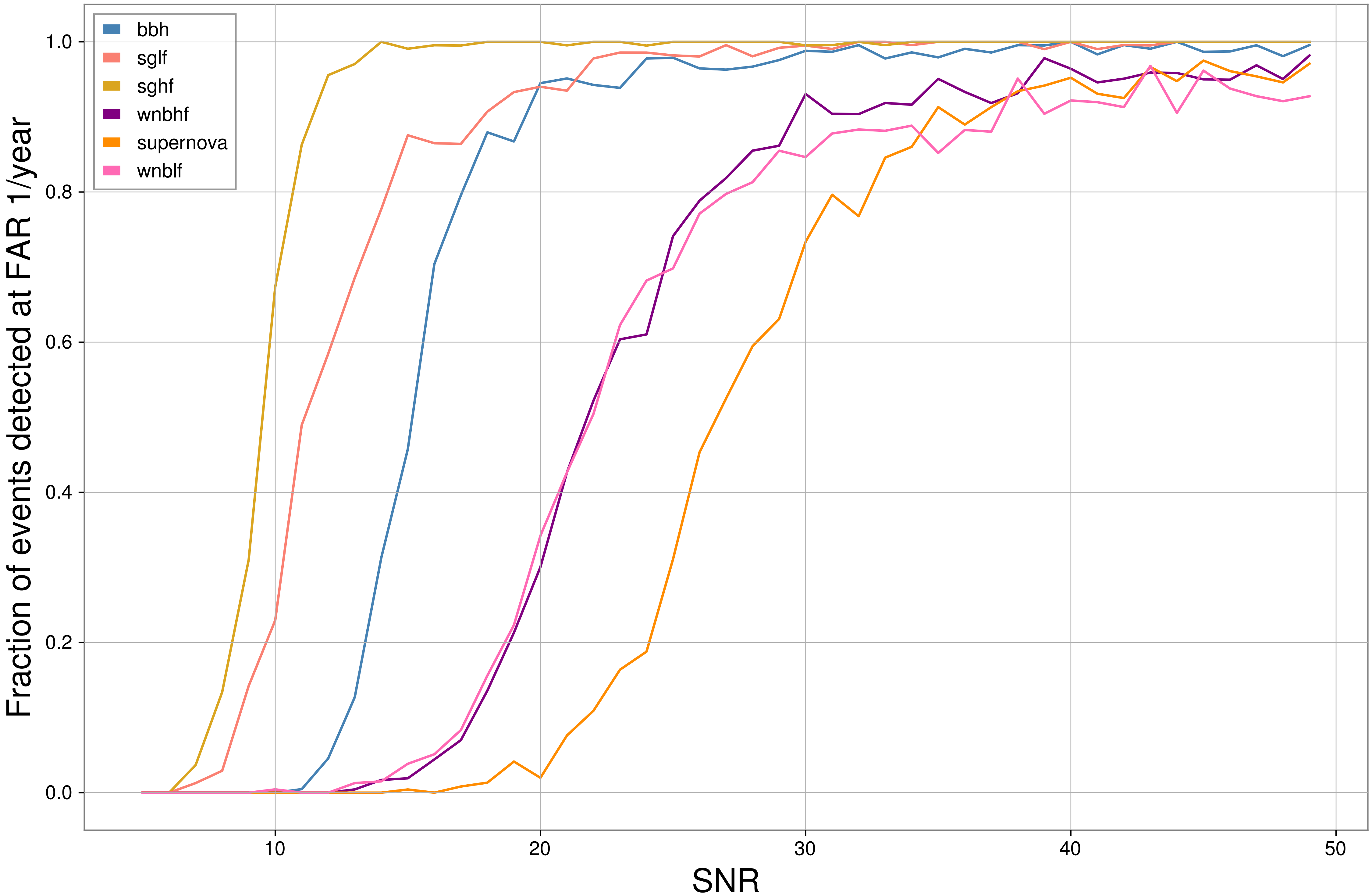


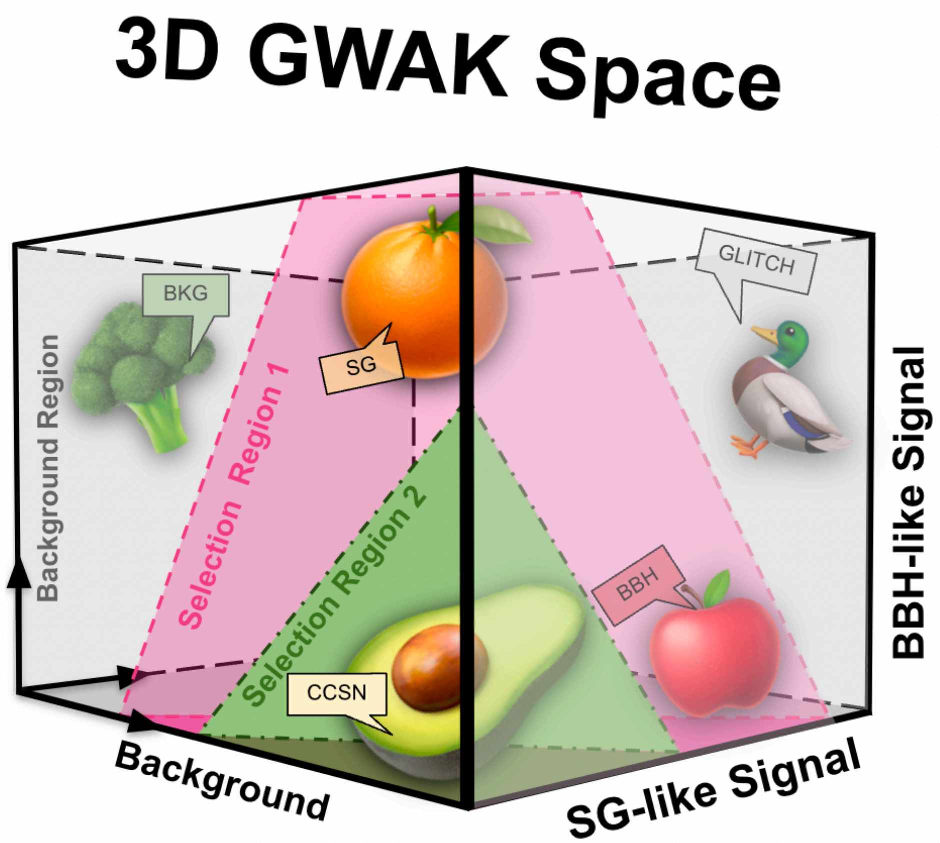
arXiv2309.11537

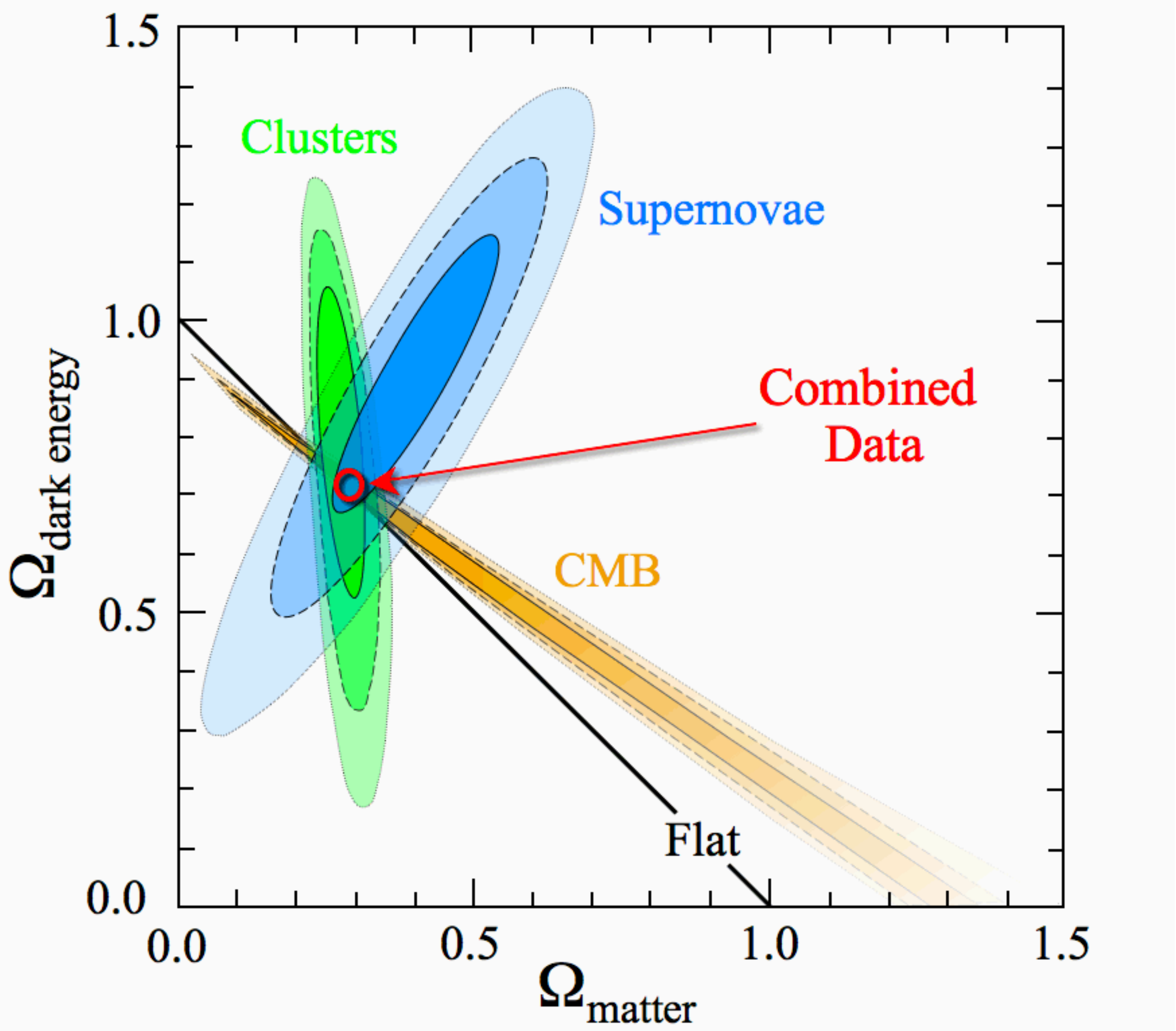
The final metric as a function of SNR for GWAK axes training signals, BBH, SG 64-512 Hz, SG 512-1024 Hz and for potential anomalies, WNB 40-400 Hz, WNB 400-1000 Hz, and Supernova



ARXIV2309.11537

- WE PRESENT A NEW SEMI-SUPERVISED APPROACH TO GRAVITATIONAL-WAVE ANOMALY DETECTION
- OUR METHOD CONSTRUCTS A 12-DIMENSIONAL EMBEDDED SPACE (GWAK SPACE) THAT IS CUT WITH A HYPERPLANE
- WE FIND THAT THE GWAK METHOD PROVIDES GOOD DISCRIMINATION POWER OVER CORRELATION METRICS



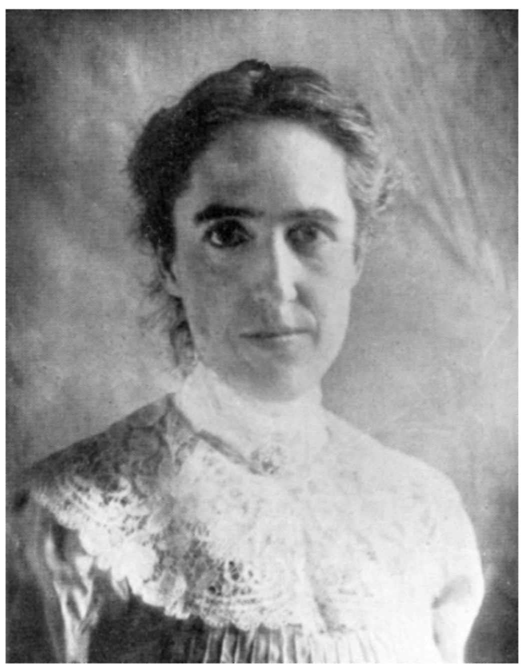



# Lecture 3: Minimizing


# Cambridge, Ma 1908

# An Observation at Harvard

Henrietta Leavitt, Harvard



Star Magnitude

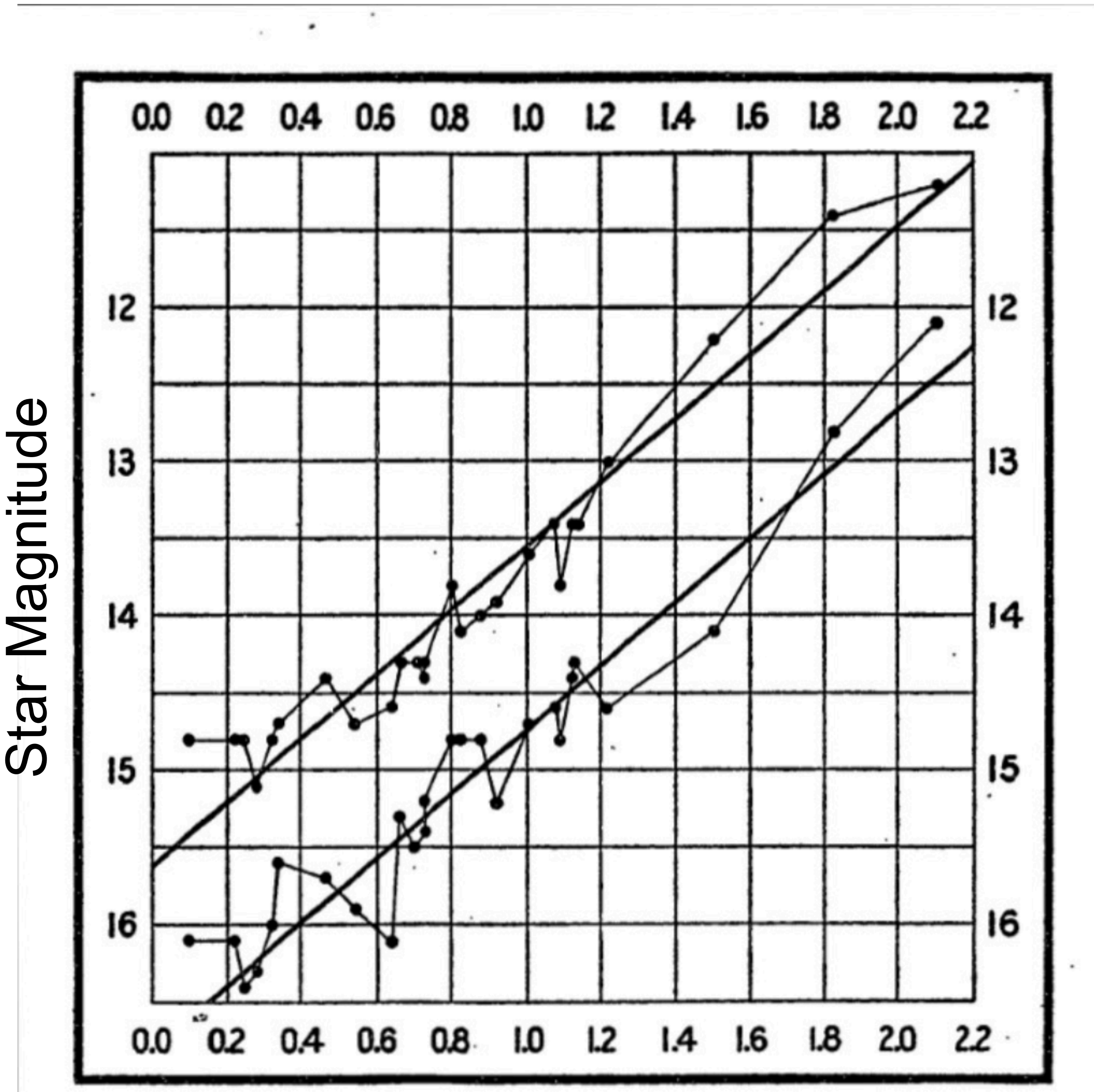
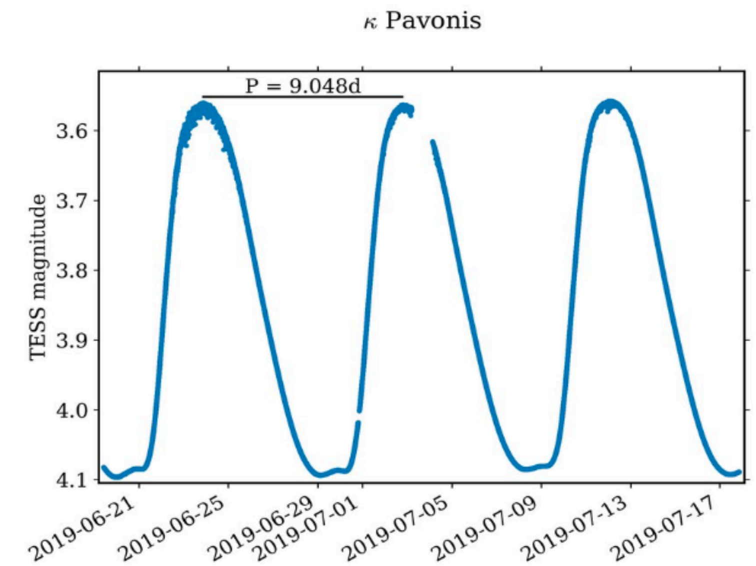



Fig. 2.

Log(Period)

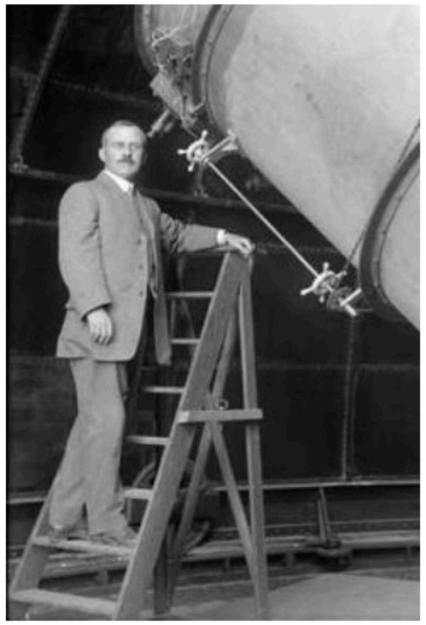


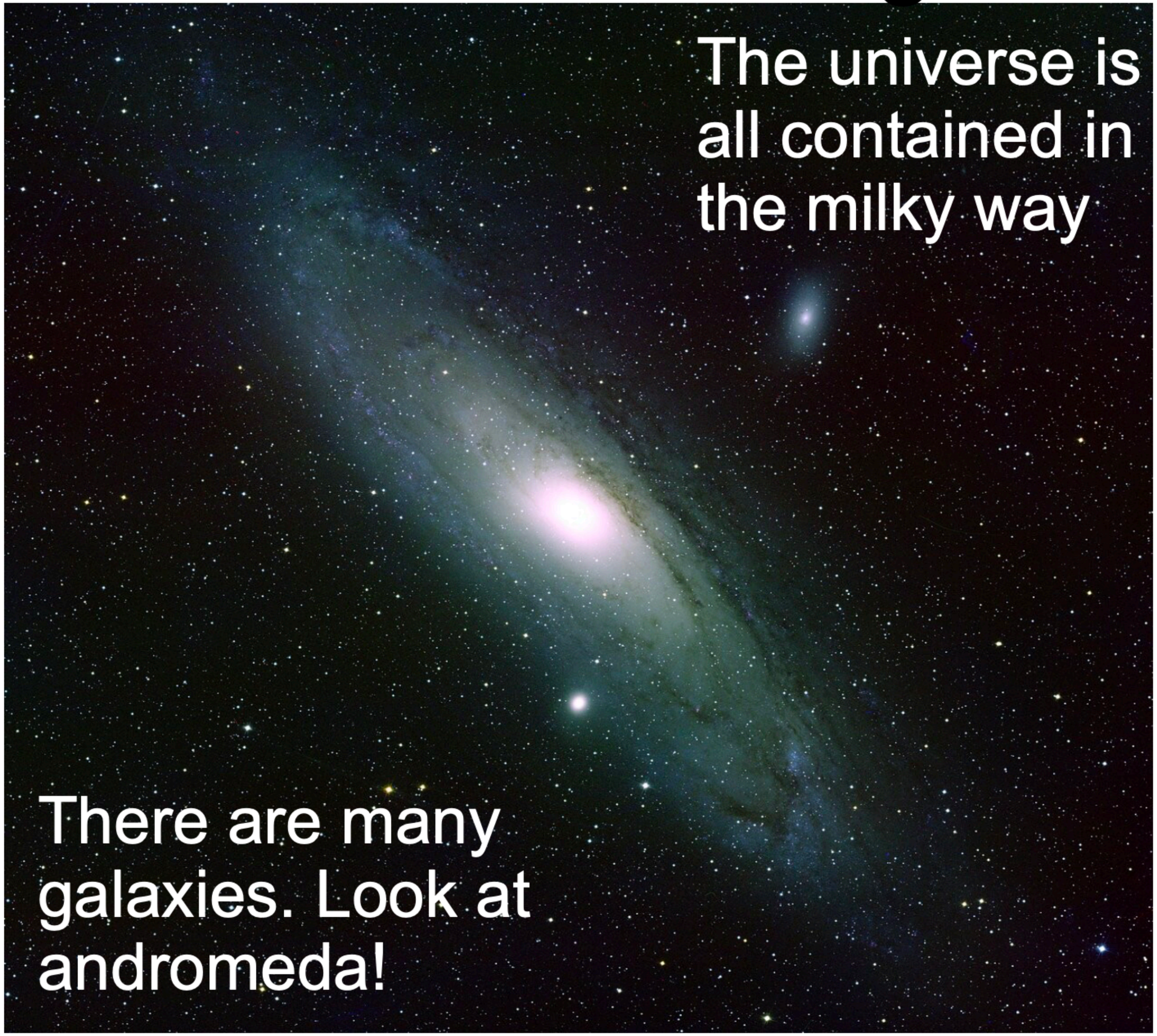

Cepheid variable stars oscillate in amplitude

# Washington, DC 1920

# The great debate

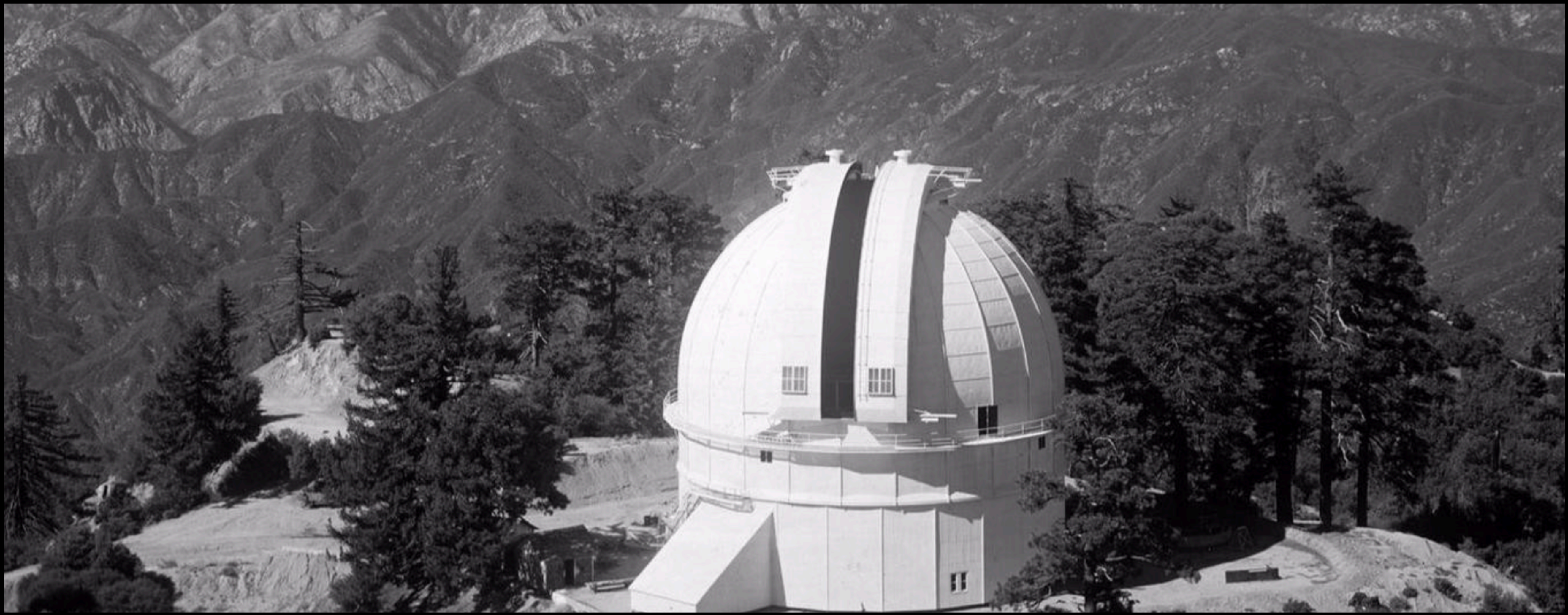
Heber Curtis, Lick (UC)






Harlow Shapely, Harvard

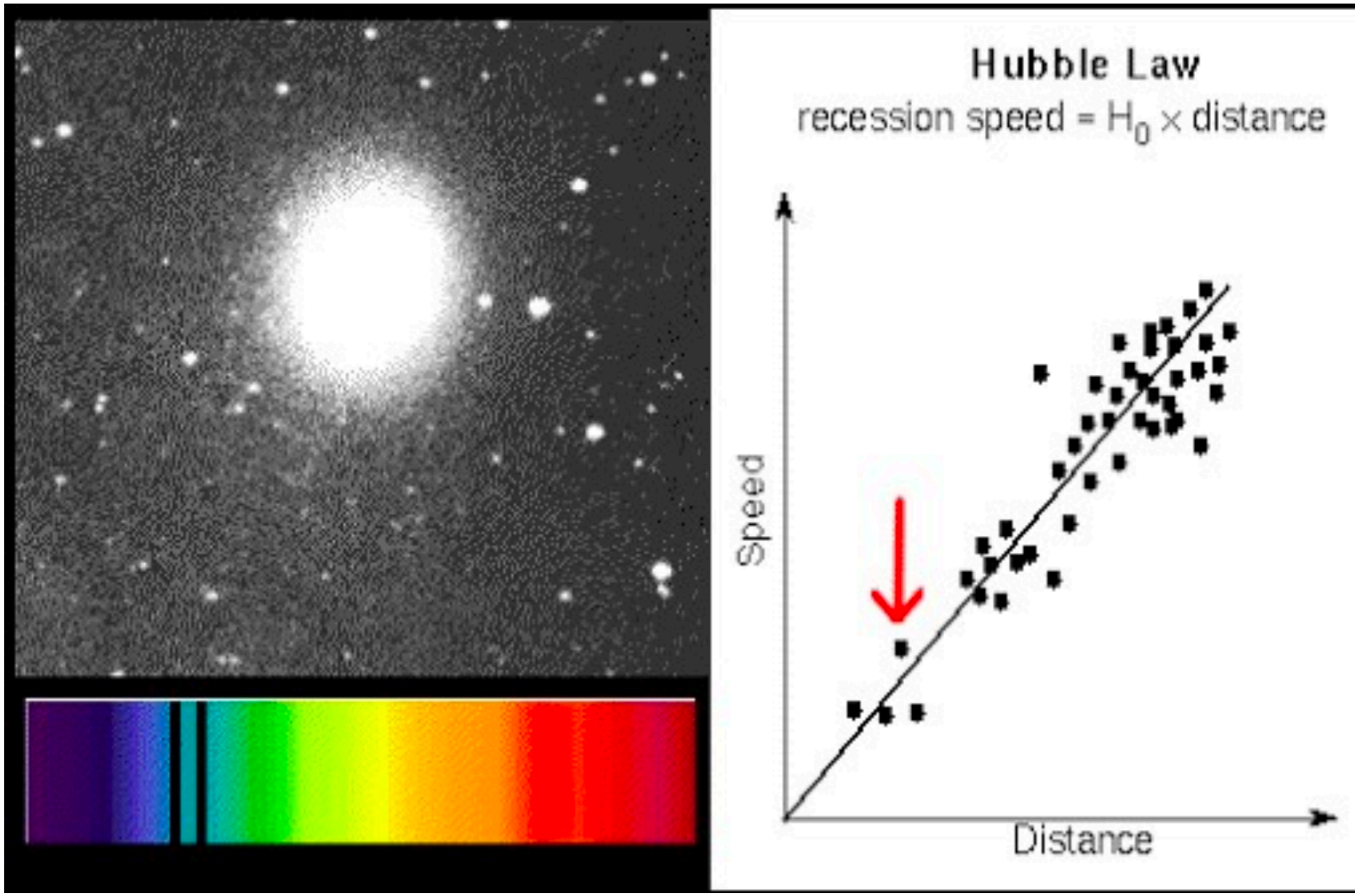
## How big is the universe



Mt. Wilson, CA 1927

# Andromeda is far away

Edwin Hubble, Caltech

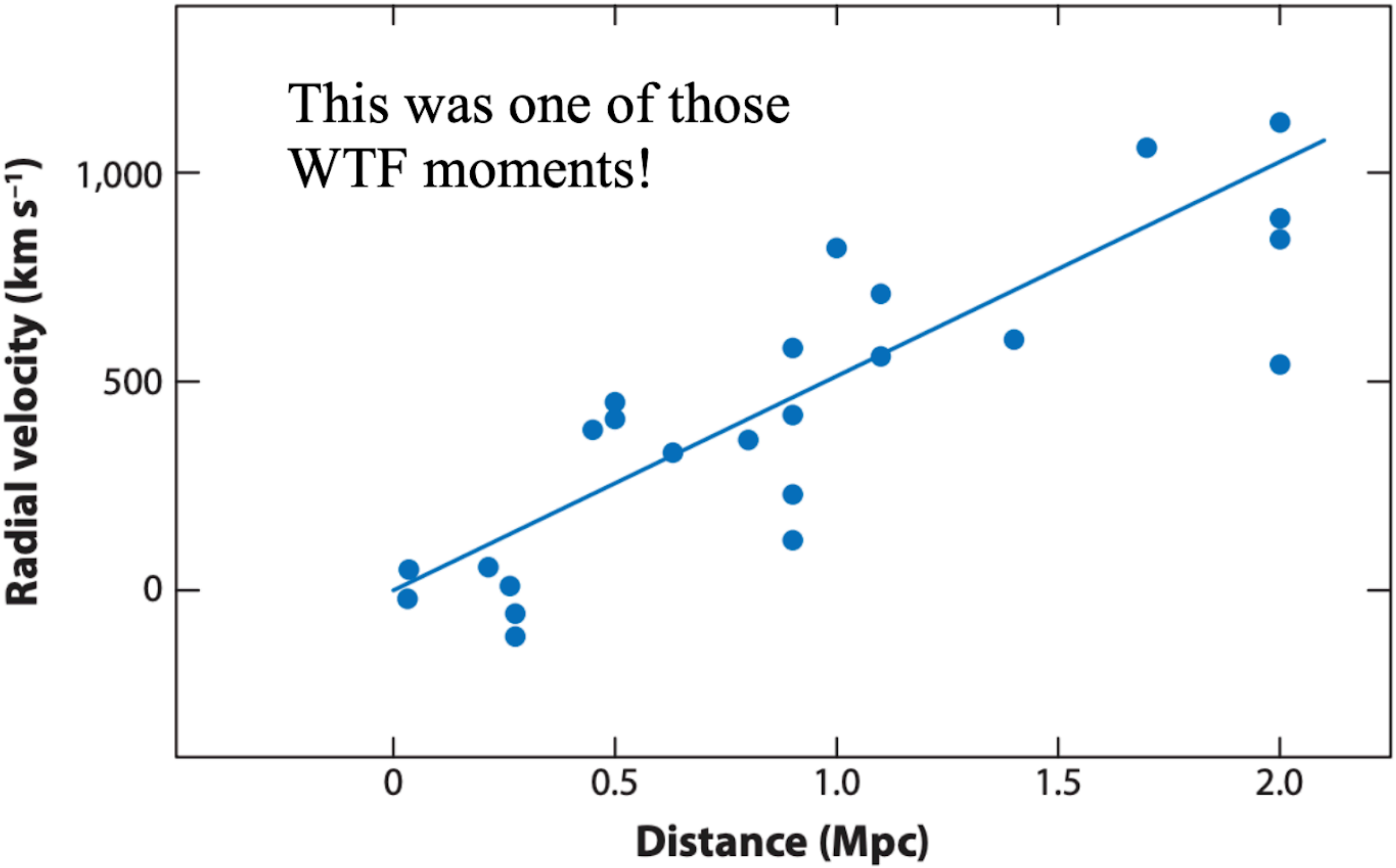



With Cepheid tech
Hubble showed Andromeda was far away

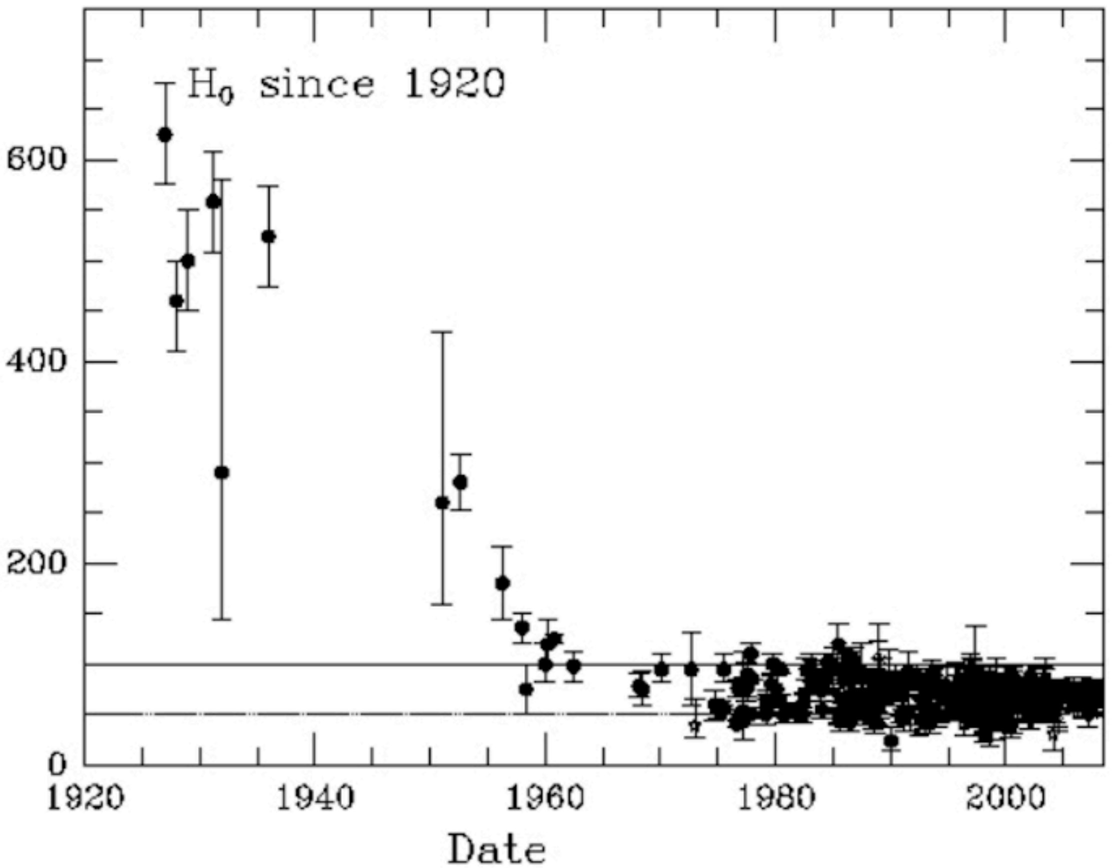
Also

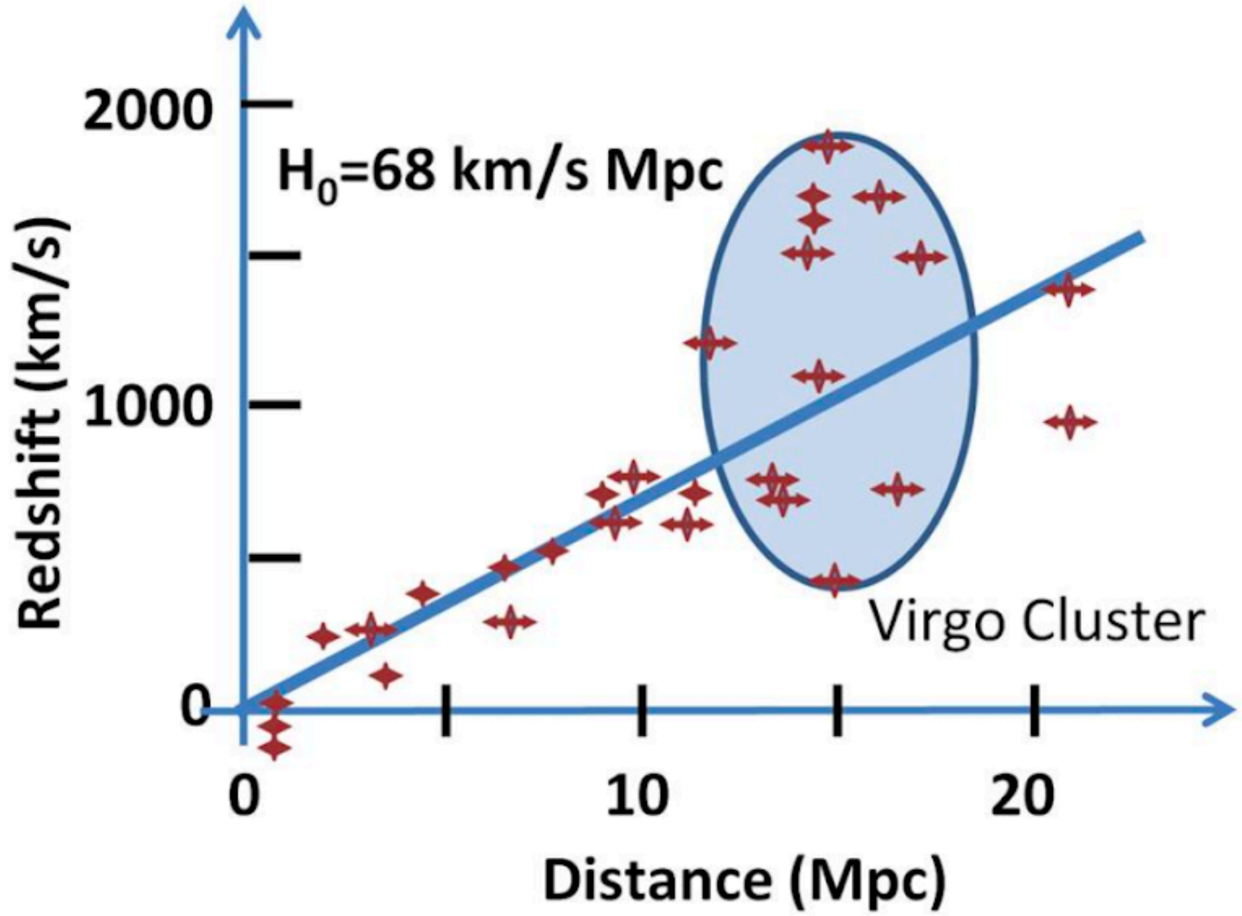
# Hubble's Expansion

All of the nearby galaxies are expanding



# Controversy of Hubble's Measurement

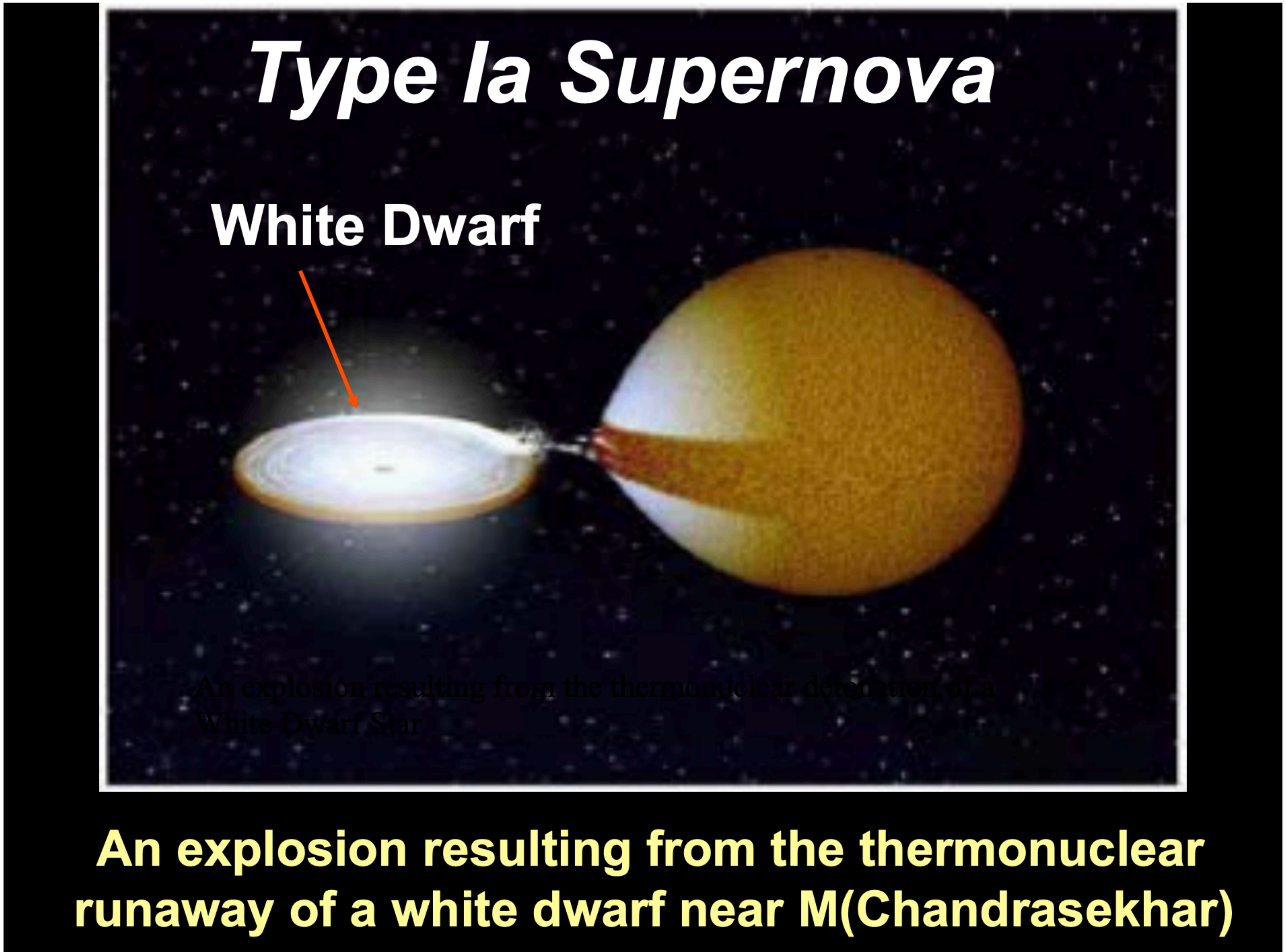



Hubble was waaaay off

Factoring in galaxy clusters and other effects are needed

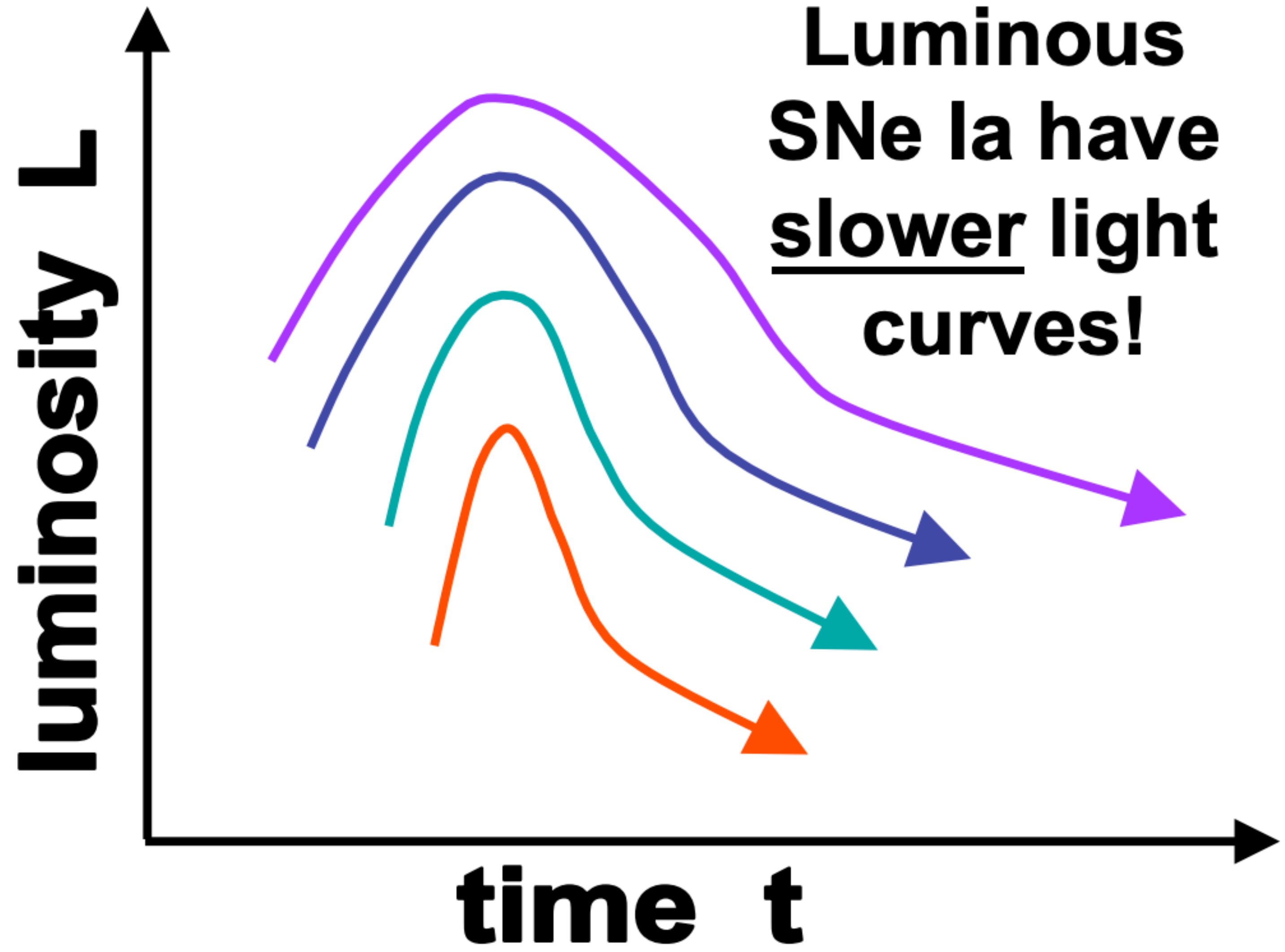
# Nowadays


Type Ia Supernova
White Dwarf
An explosion resulting from the thermonuclear detonation of a White Dwarf Star
An explosion resulting from the thermonuclear runaway of a white dwarf near M(Chandrasekhar)

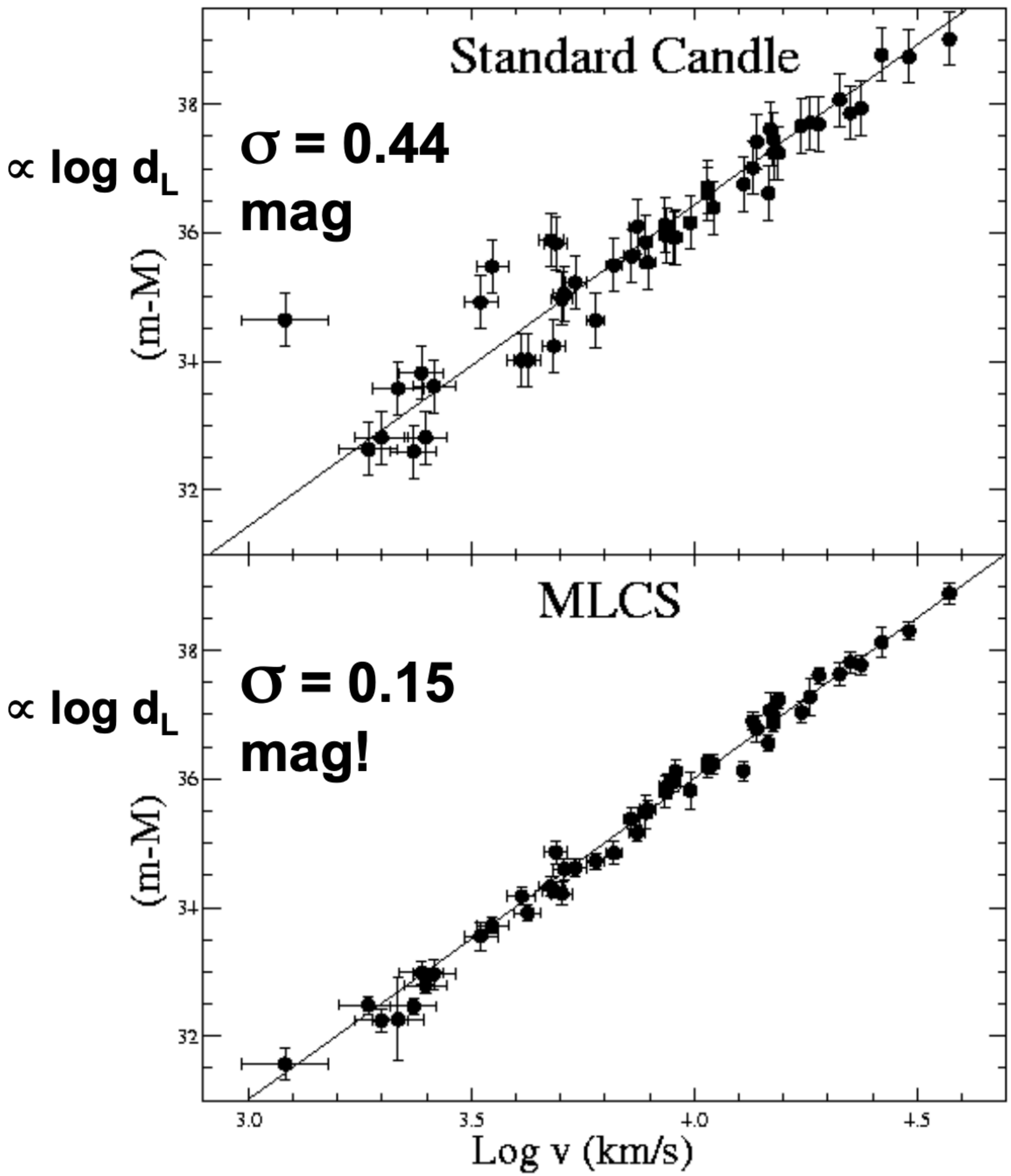
# Calibrating the Nearly Standard Candle

- Phillips (1993), Riess + (1995), Hamuy+ (1995): established L vs. light-curve shape correlation with ~ 10 nearby SNe Ia
- Use it to standardize other SNe Ia
  - Measured colors give reddening and extinction
  - Accurately calibrate individual SNe!

Absolute light curves of SN Ia in galaxies of known distance

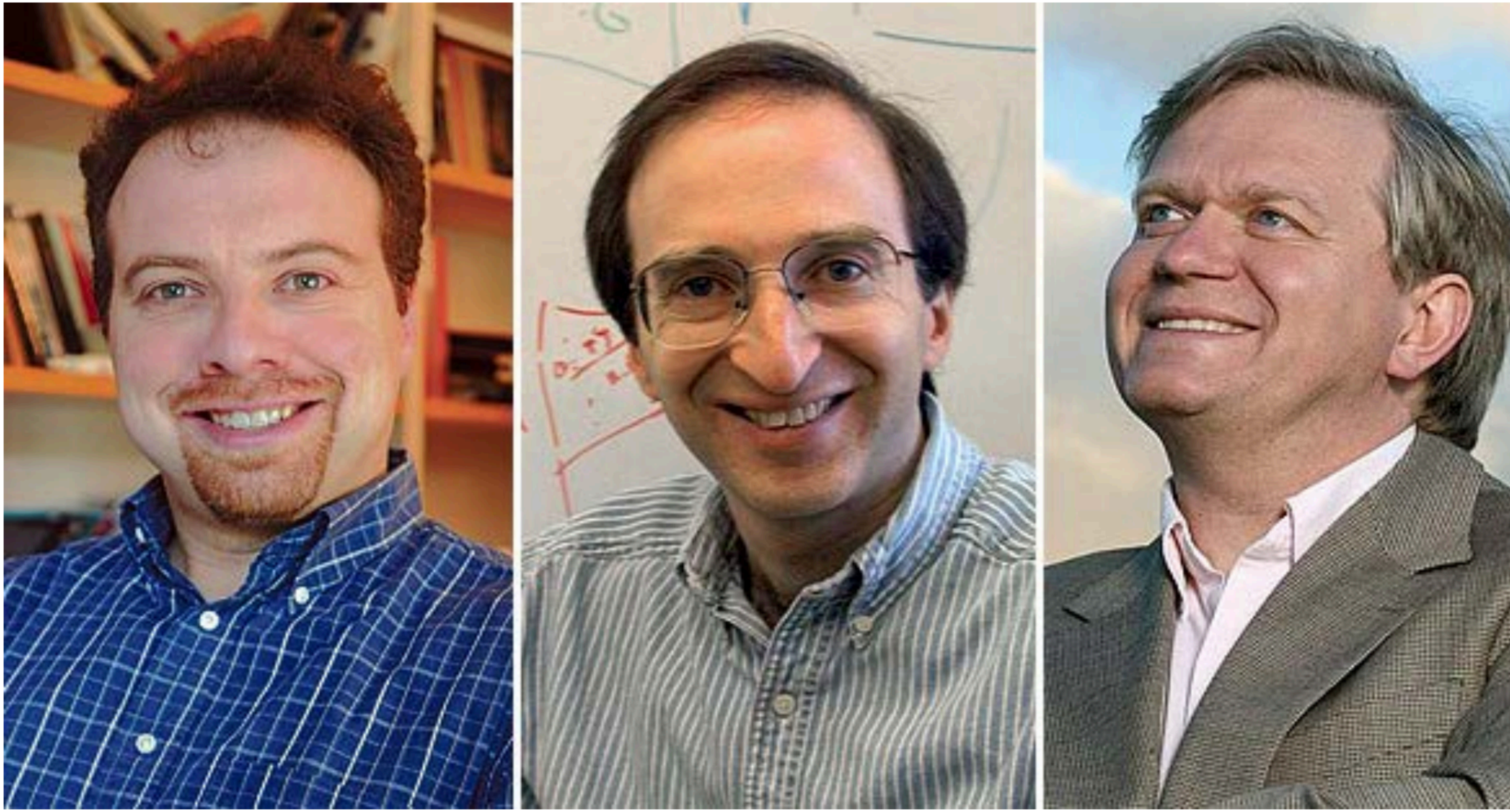



**MLCS: Multi-color Light-Curve Shape (Riess et al.)**




Calibration

# Studies of Universe's Expansion Win Physics Nobel



Johns Hopkins University; University Of California At Berkeley; Australian National Univers

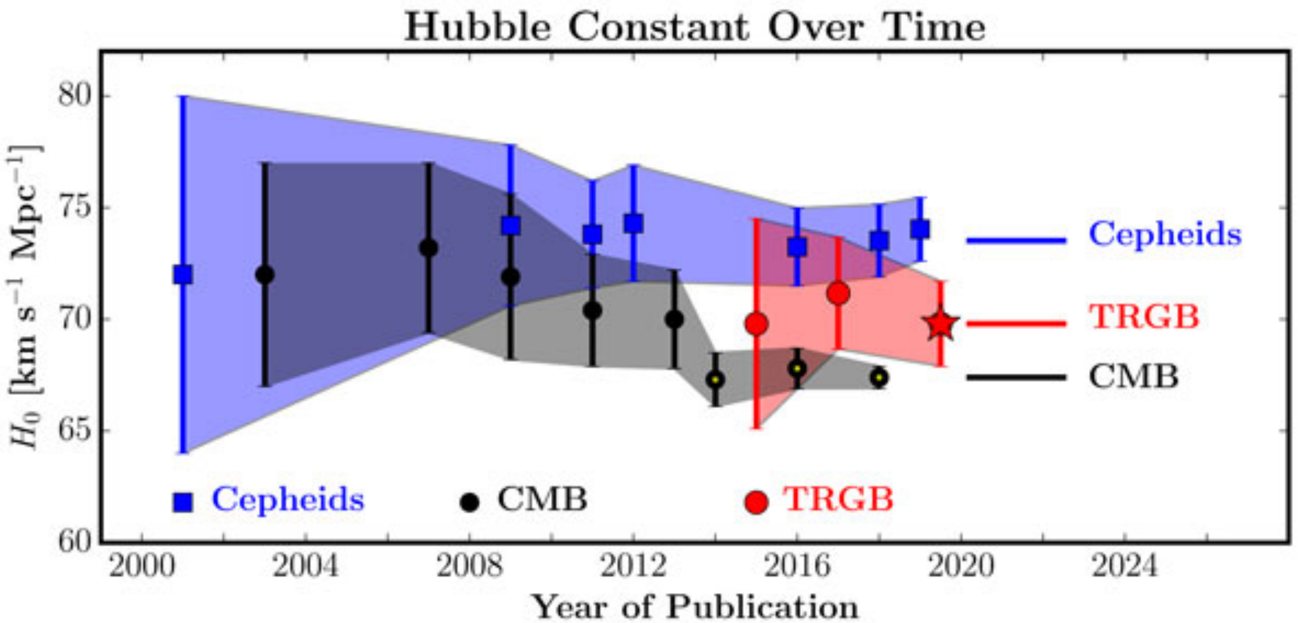
From left, Adam Riess, Saul Perlmutter and Brian Schmidt shared the Nobel Prize in physics awarded Tuesday.

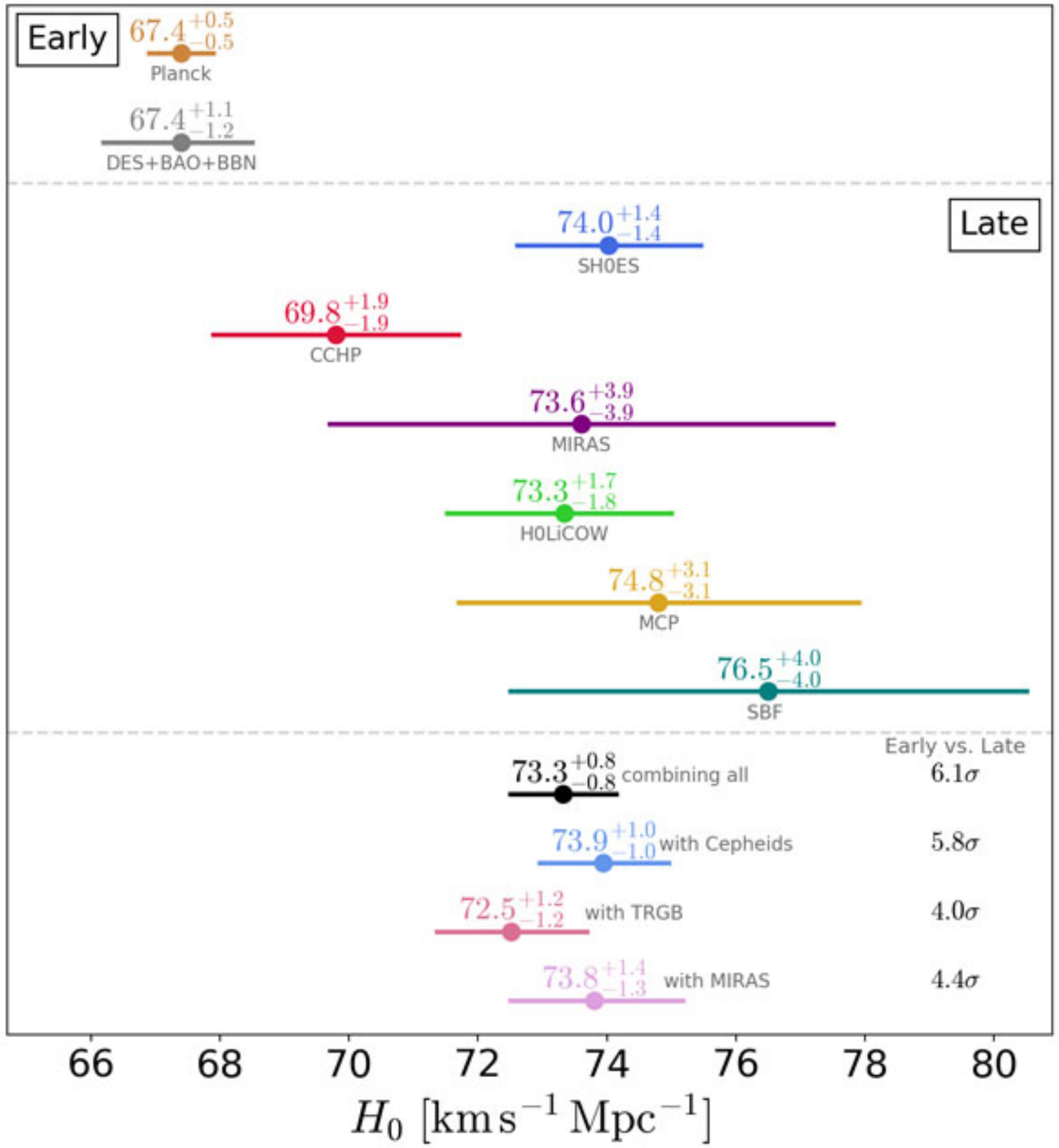
By DENNIS OVERBYE
Published: October 4, 2011

## 2011 Nobel Prize in Physics

# However there appears to be tension

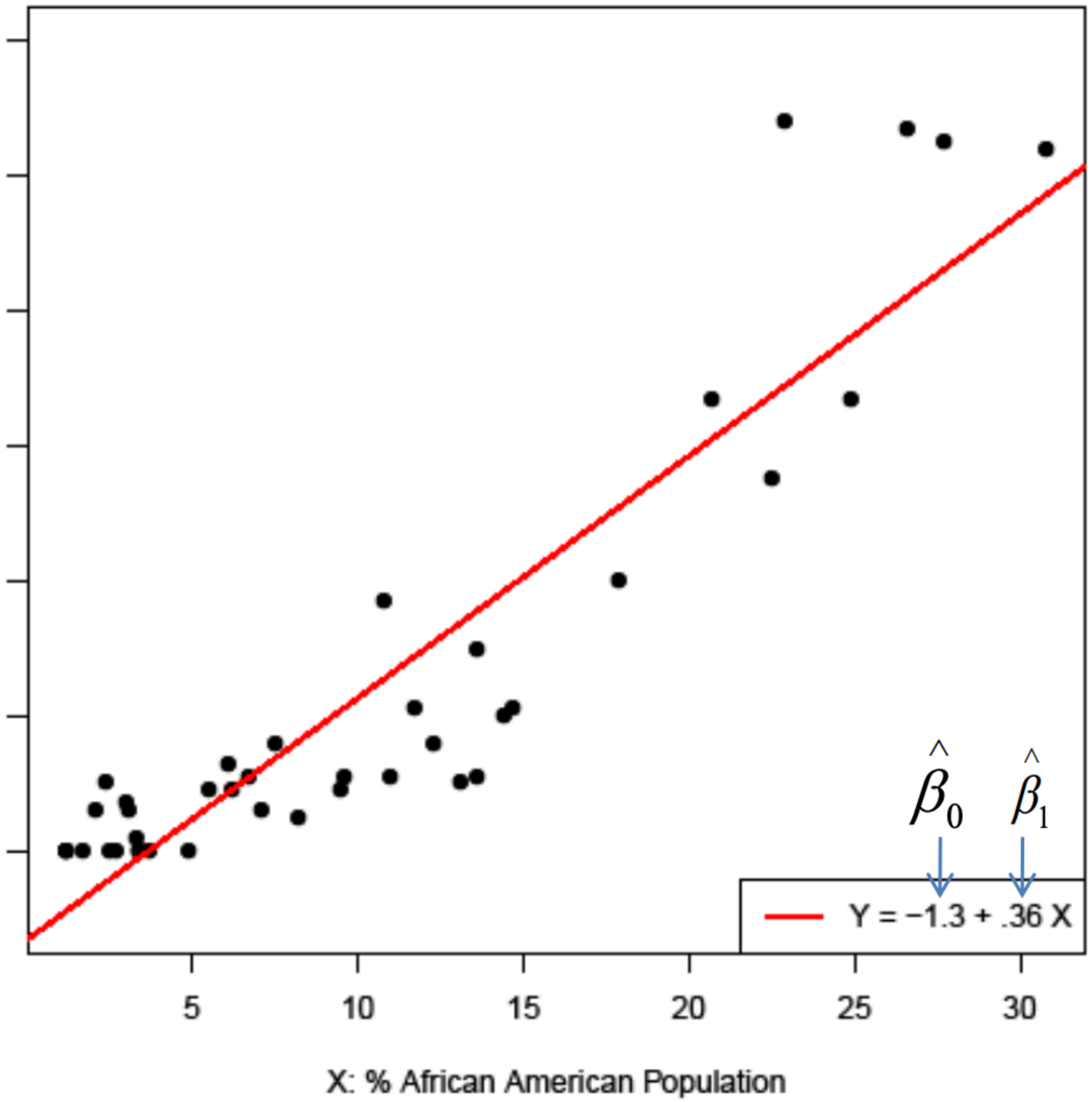
- Current Supernovae+Cepheid measurements
  - Indicate a larger value of Hubble's constant
- CMB indicates a smaller value
- Lets do the fit ourselves!

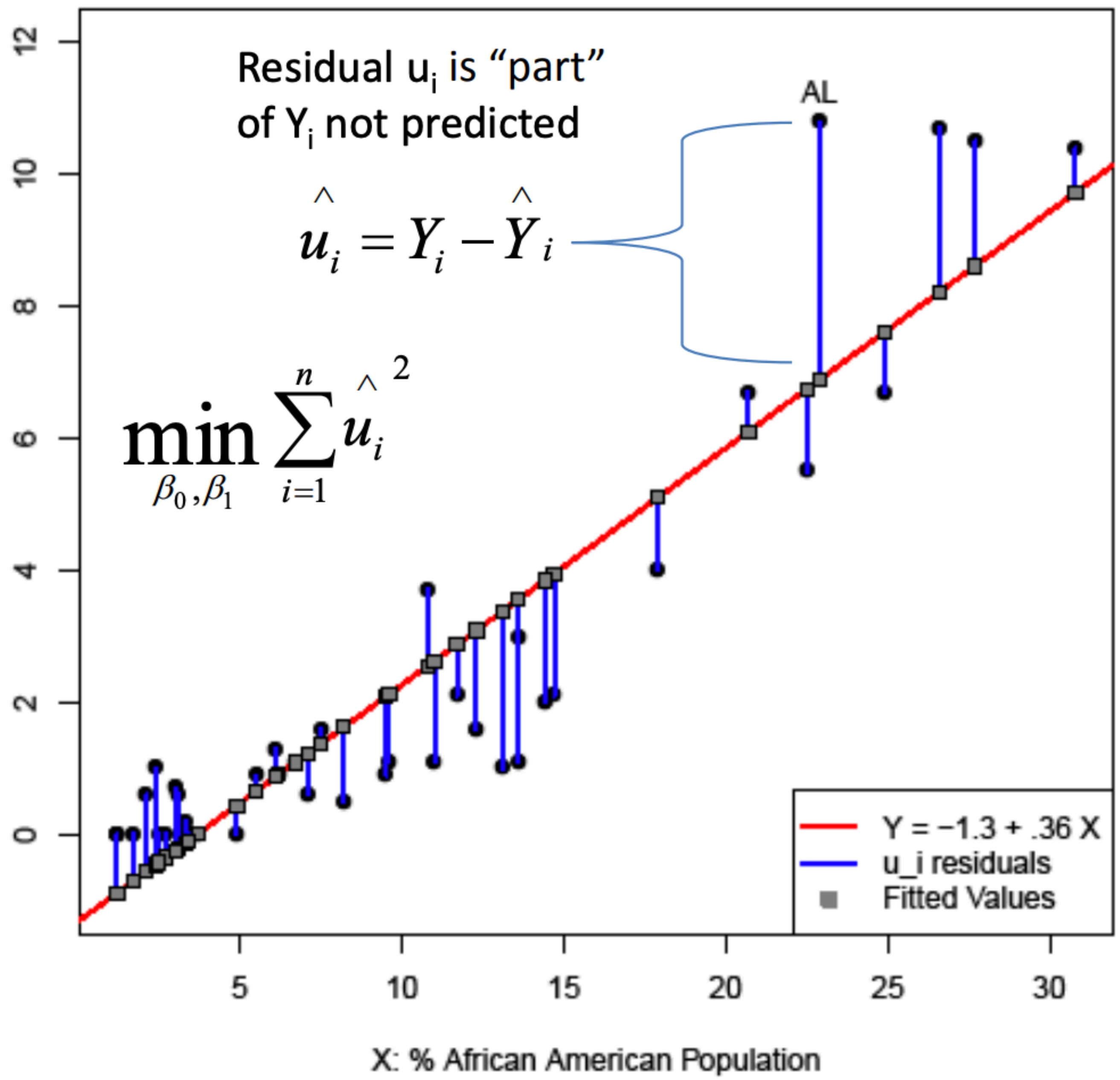


# Linear Regression from scratch

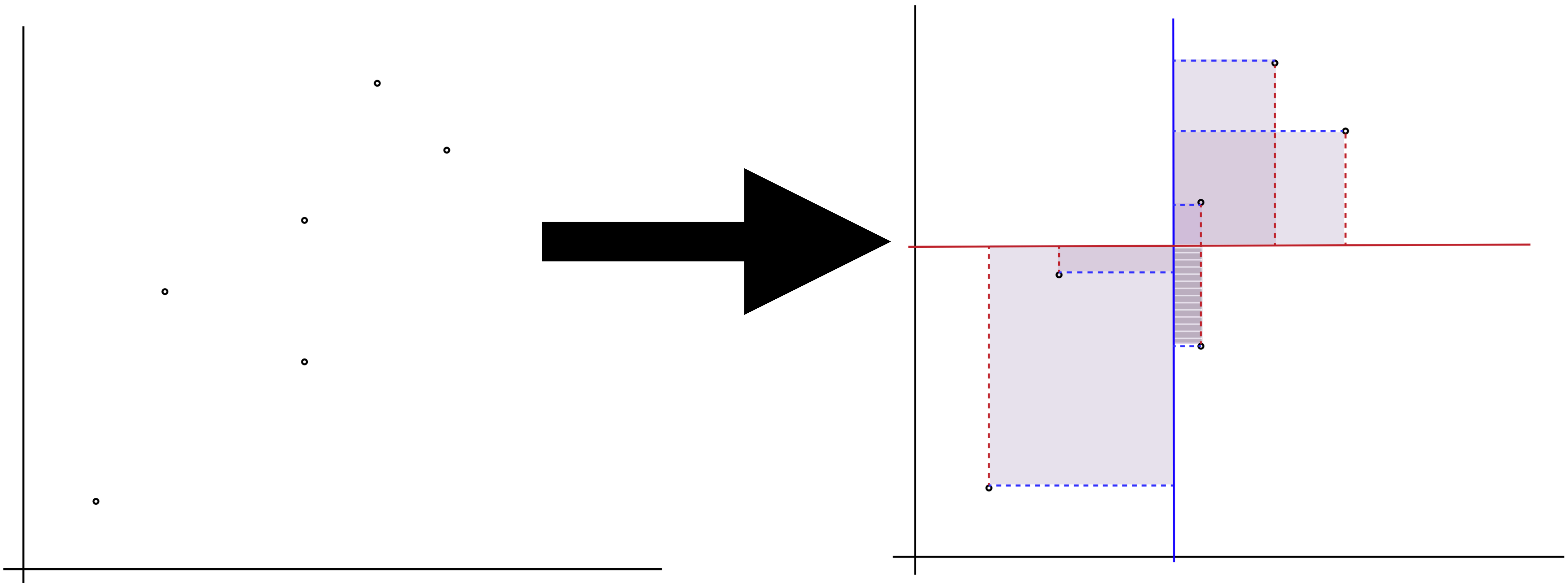
# What do we want to mininize?



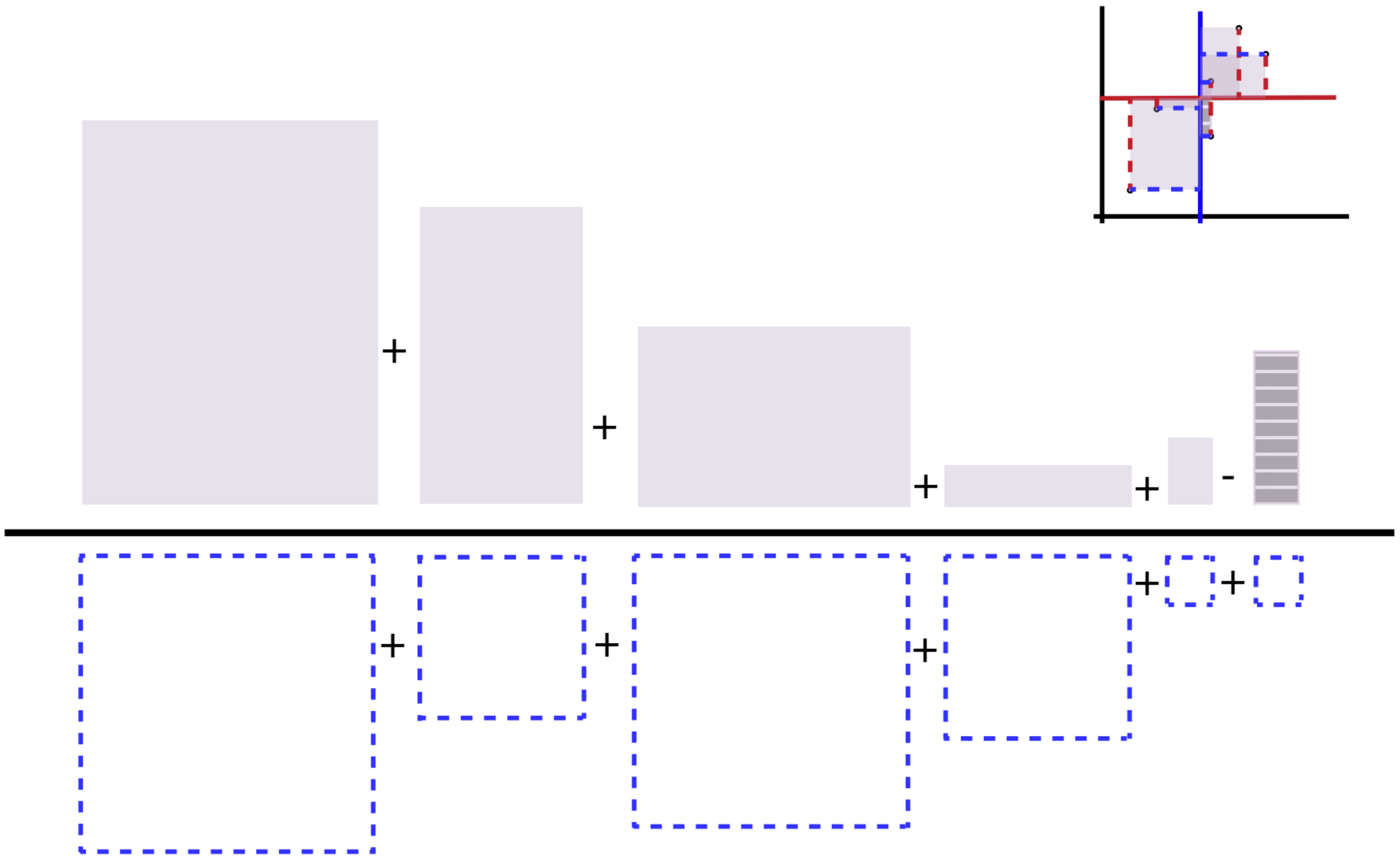




https://scholar.princeton.edu/sites/default/files/bstewart/files/lecture5handout.pdf

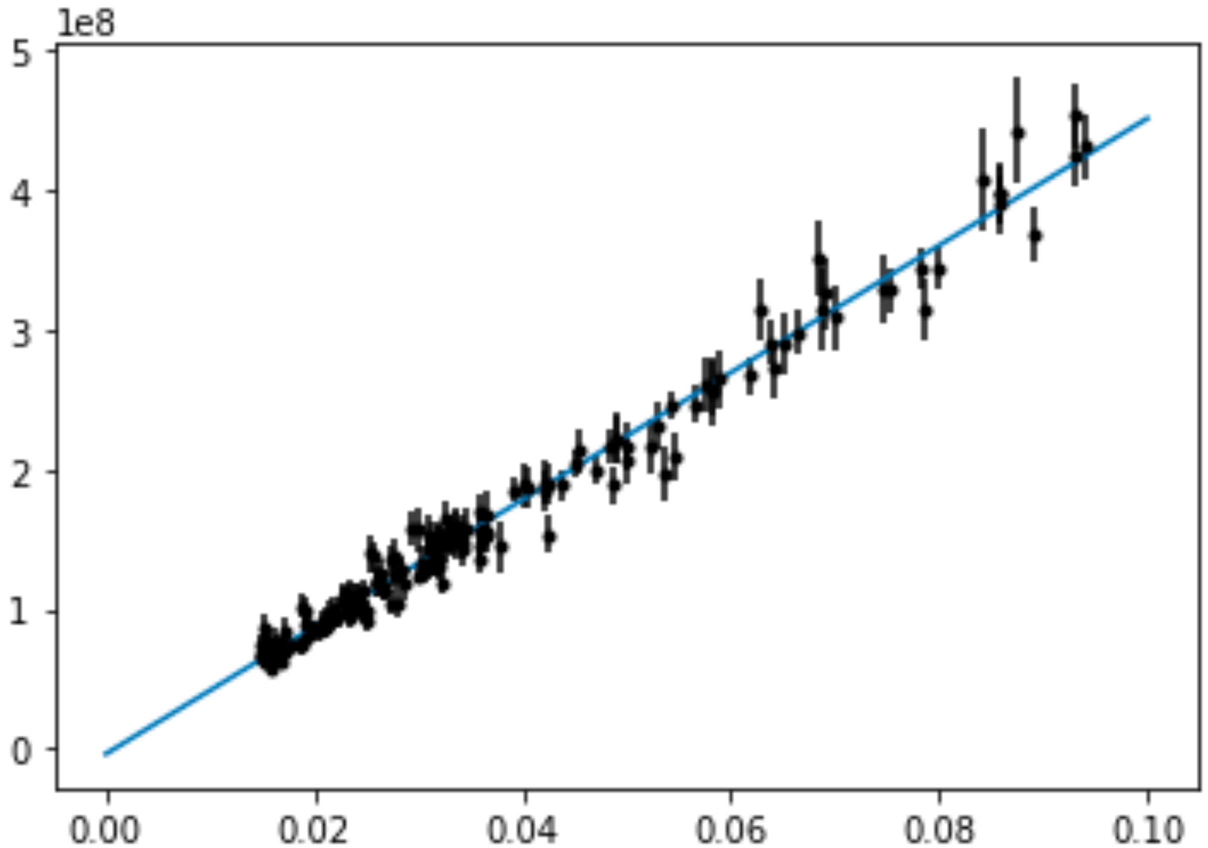
# Visualizing Derivation



# Visualizing Derivation



# Hubble's Measurement




Distance

Redshift

- A = Distance (pc)/redshift = (Mpc/$10^6$)/(c/v)/Redshift
- Redshift = v/c => Redshift = v / c($3\text{x}10^6$ km/s) = (v in km/s)/$3\text{x}10^6$ km/s
- A = 1/$10^6$ (Distance in Mpc) / ( v in Km/s / $3\text{x}10^6$ km/s)) = 1/($3\text{x}10^6$ x$10^6$) (Distance in Mpc) / (velocity in km/s)
- Hubble's Constant: $H_0$ = 67 km/s/Mpc which is velocity in km/s over distance in Mpc
- $H_0$ = ($3\text{x}10^6$ x$10^6$)/A With error propagation $\rightarrow \sigma_h = (3 \times 10^{12}/A^2)\sigma_A$

# Regression Forms

$$A = \frac{Cov(X,Y)}{Var(X)} \text{ or } \frac{\sum xy}{\sum x^2} \text{ or } \frac{\sum_{i=1}^{n}(X_iY_i) - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n}(X_i)^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}} \text{ or } \frac{\sum_{i=1}^{n}(X_iY_i) - n\overline{X}\,\overline{Y}}{\sum_{i=1}^{n}\left(X_i^2\right) - n\overline{X}^2} \text{ or } \frac{\frac{1}{n}\sum_{i=1}^{n}(X_iY_i) - \overline{X}\,\overline{Y}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2\right) - \overline{X}^2} \text{ or } \frac{\overline{(XY)} - \overline{X}\,\overline{Y}}{\overline{\left(X_i^2\right)} - \overline{X}^2}$$

- There are a lot of different ways to write the regression

# More Regression Forms

$$\sum x^2 = SS_x = (n-1)Var(X) = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \sum_{i=1}^{n}\left(X_i^2\right) - n\overline{X}^2$$

$$\sum y^2 = SS_y = (n-1)Var(Y) = \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 = \sum_{i=1}^{n}\left(Y_i^2\right) - n\overline{Y}^2$$

$$\sum xy = S_{xy} = (n-1)Cov(X,Y) = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right) = \sum_{i=1}^{n}\left(X_i Y_i\right) - n\overline{X}\,\overline{Y}$$

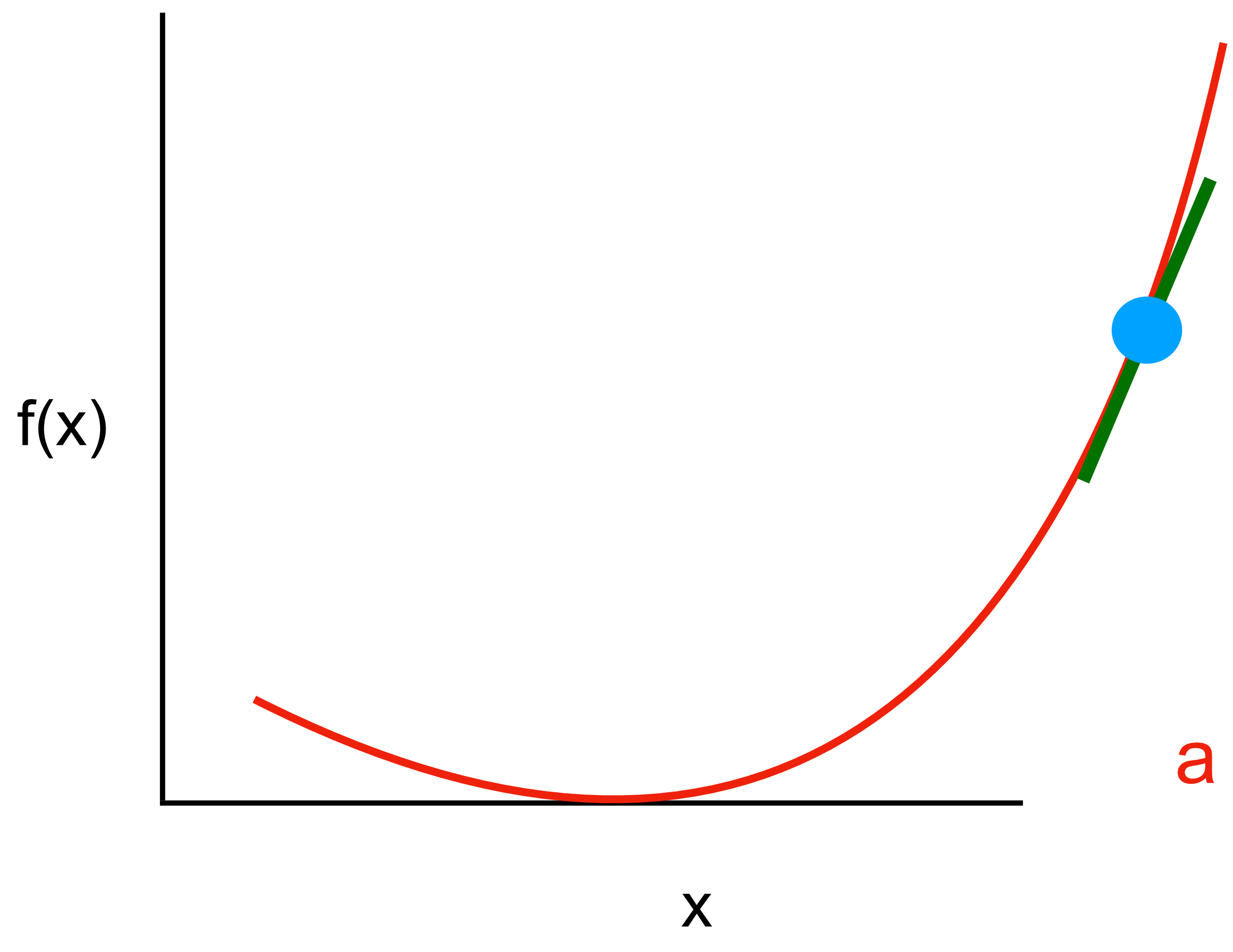
- Here we refer to SS as the sum of the squares

# Backup

# Gradient Descent

- The ideas are very similar to real life
- How do we get to the minimum as fast as possible

# Gradient



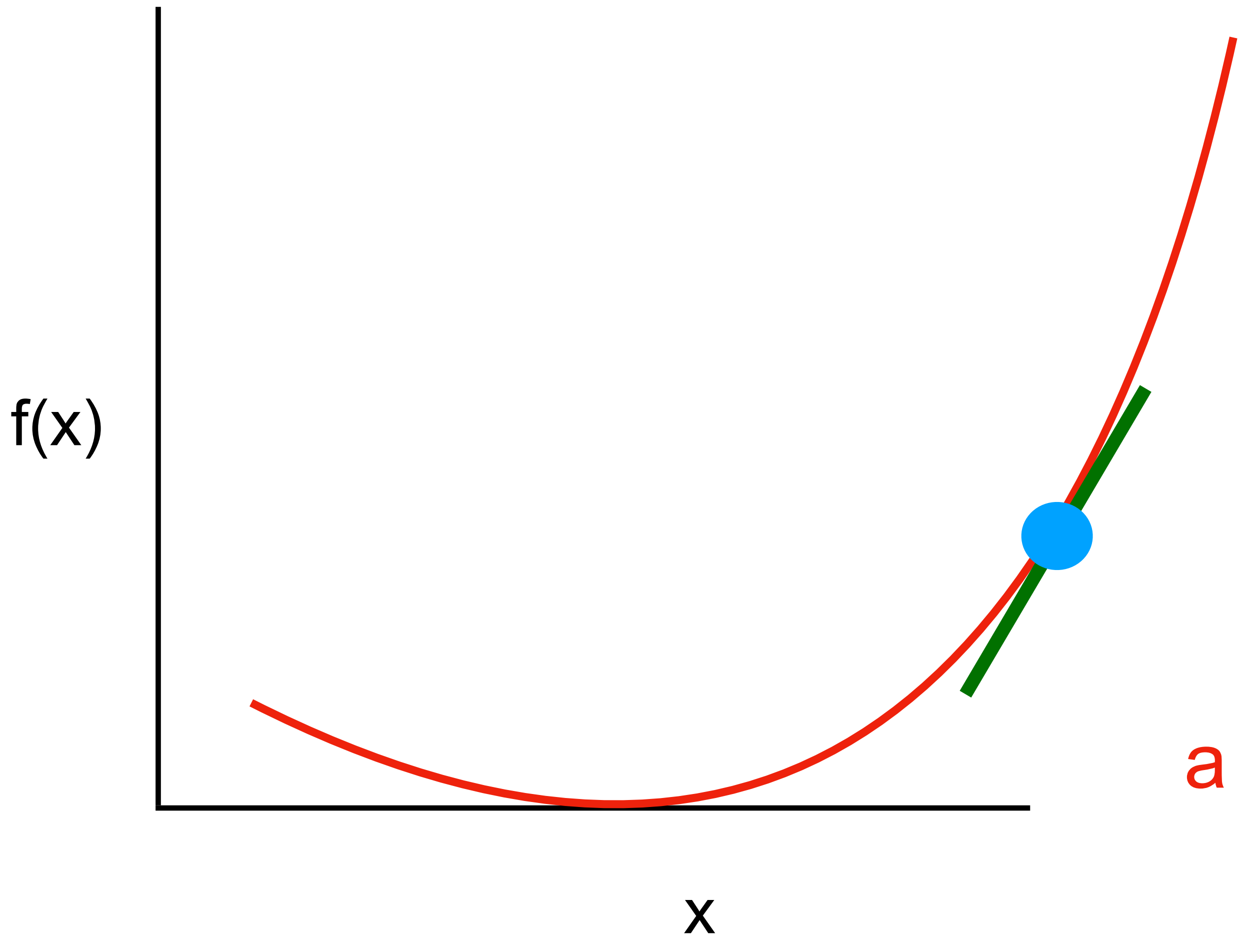

Simplest approach let
a ball roll down a potential

$$F = -\nabla U$$

$$\vec{v} = \overrightarrow{v_0} - \delta t \nabla f$$

$$\vec{x} = \overrightarrow{x_0} + \delta t \vec{v} - \frac{1}{2}(\delta t)^2 \nabla f$$ +Dampen

# Gradient



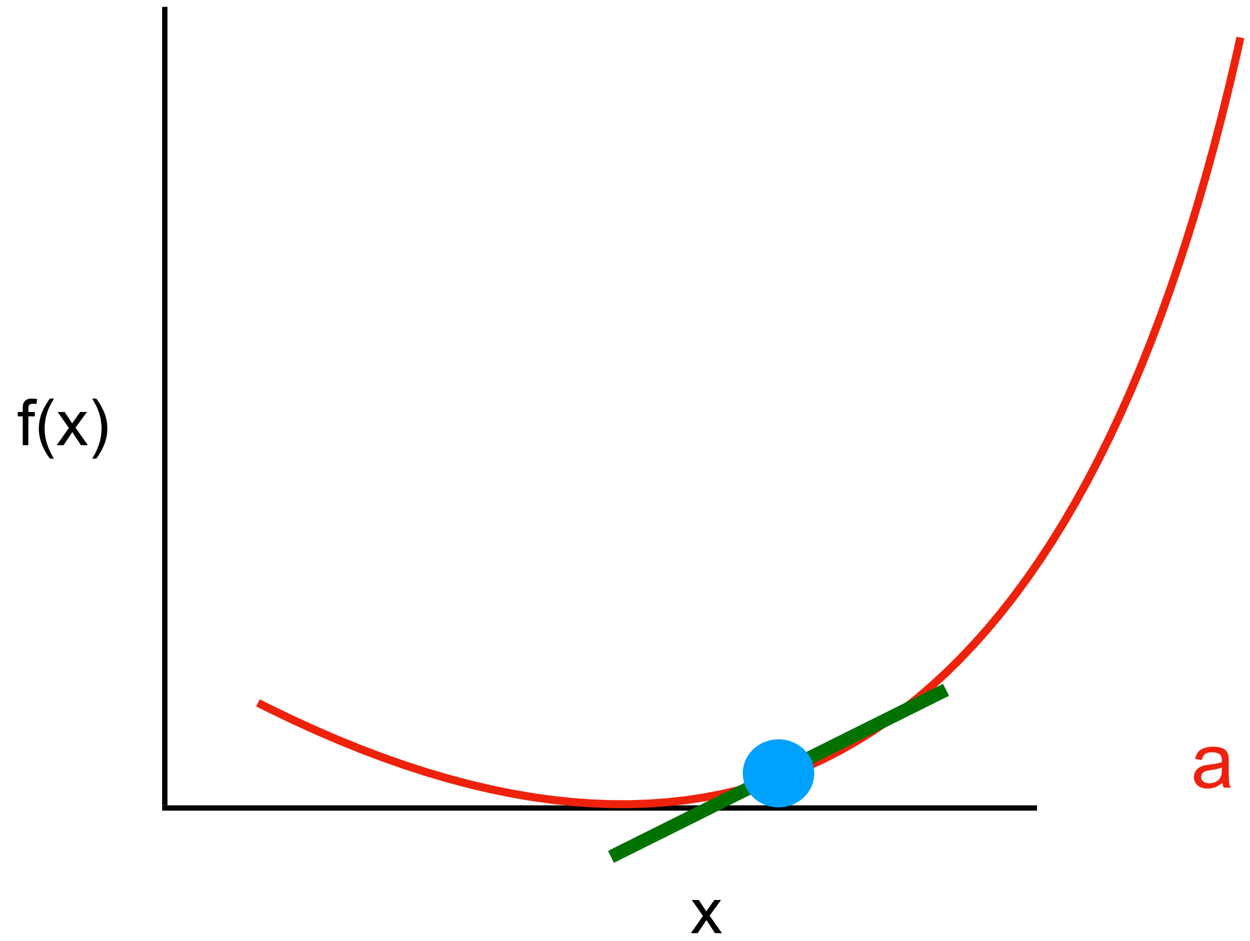

Simplest approach let
a ball roll down a potential

$$F = -\nabla U$$

$$\vec{v} = \overrightarrow{v_0} - \delta t \nabla f$$

$$\vec{x} = \overrightarrow{x_0} + \delta t \vec{v} - \frac{1}{2}(\delta t)^2 \nabla f$$ +Dampen

# Gradient



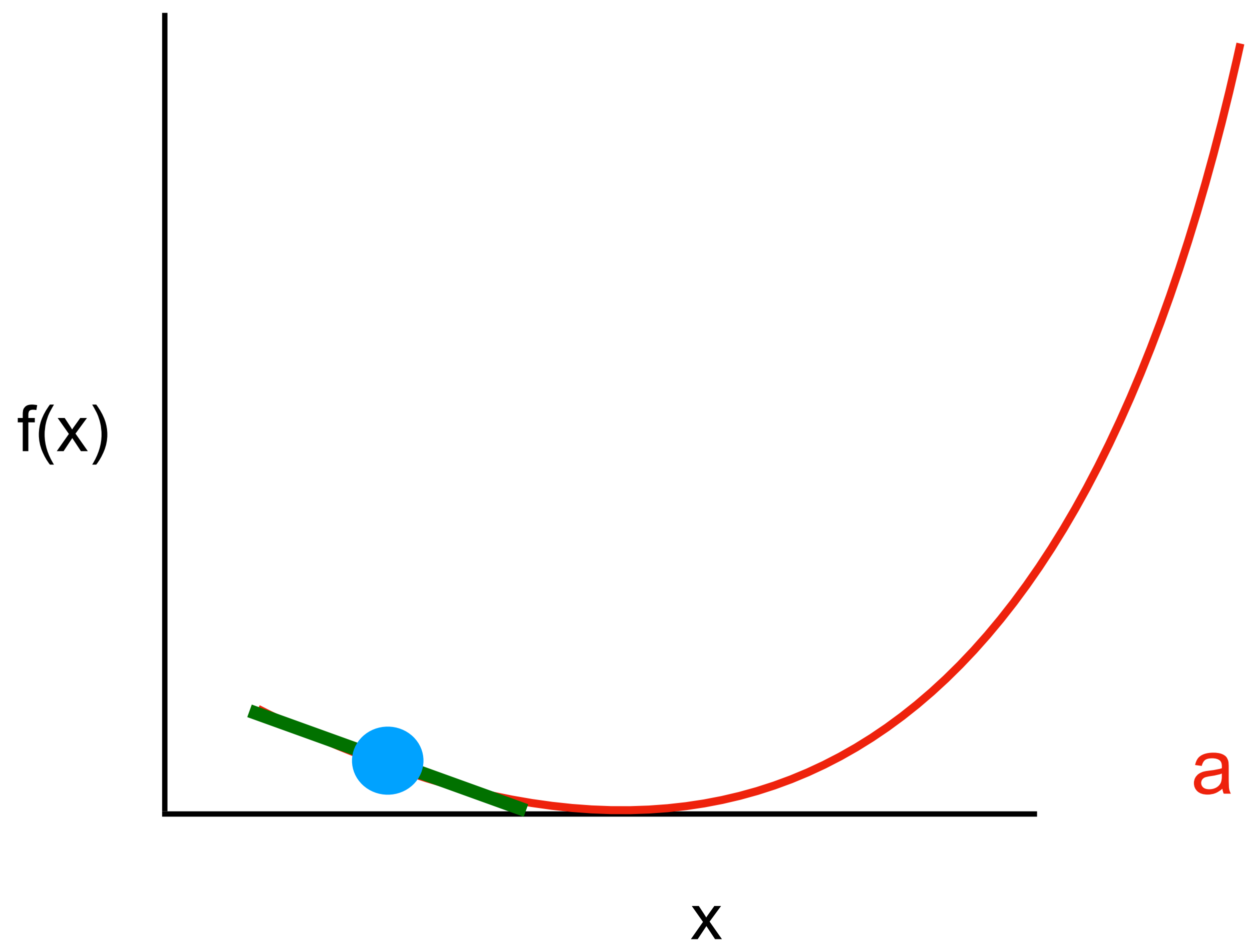

Simplest approach let
a ball roll down a potential

$$F = -\nabla U$$

$$\vec{v} = \vec{v_0} - \delta t \nabla f$$

$$\vec{x} = \vec{x_0} + \delta t \vec{v} - \frac{1}{2}(\delta t)^2 \nabla f$$ +Dampen

# Gradient



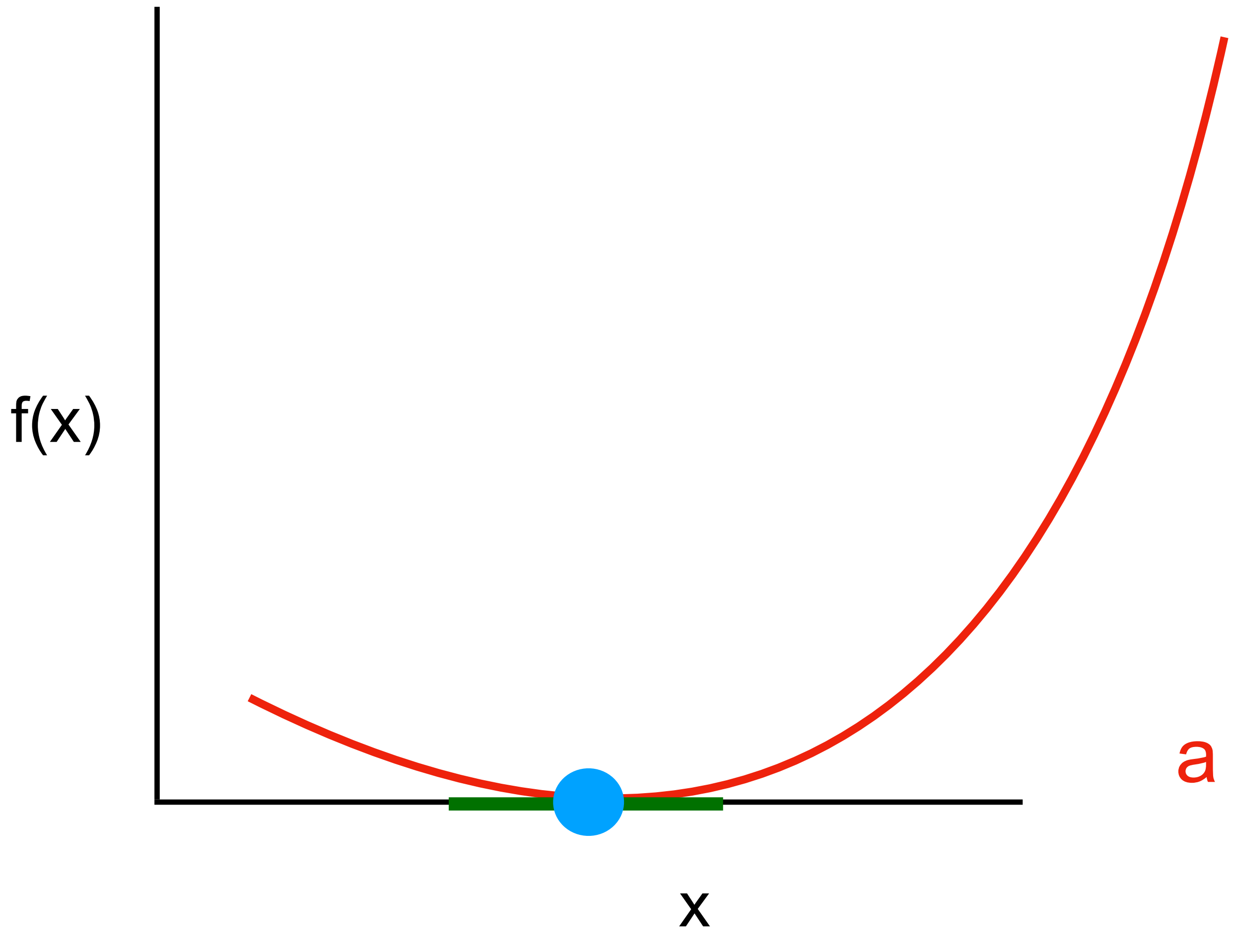

Simplest approach let
a ball roll down a potential

$$F = -\nabla U$$

$$\vec{v} = \overrightarrow{v_0} - \delta t \nabla f$$

$$\vec{x} = \overrightarrow{x_0} + \delta t \vec{v} - \frac{1}{2}(\delta t)^2 \nabla f$$ +Dampen

# Gradient




Simplest approach let
a ball roll down a potential

$$F = -\nabla U$$

$$\vec{v} = \overrightarrow{v_0} - \delta t \nabla f$$

$$\vec{x} = \overrightarrow{x_0} + \delta t \vec{v} - \frac{1}{2}(\delta t)^2 \nabla f$$ +Dampen

# Newton’s Step

$$f(x+h) = f(x) + h\frac{df}{dx}(x) + \frac{h^2}{2}\frac{d^2f}{dx^2}$$

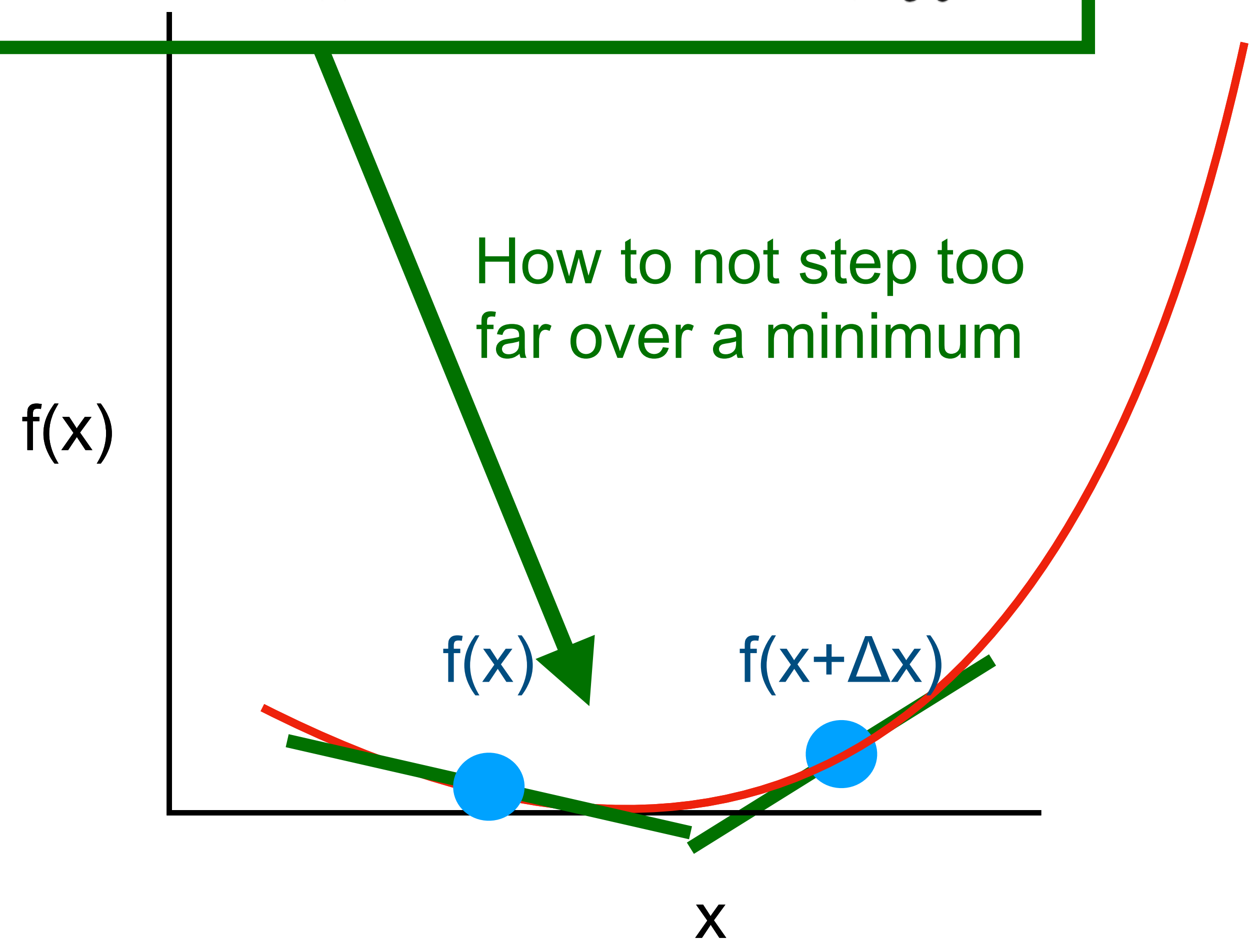
$$\boxed{\frac{f(x+h)-f(x)}{h} = 0 = \frac{df}{dx}(x) + \frac{h}{2}\frac{d^2f}{dx^2}}$$

$$0 = \frac{df}{dx}(x) + \Delta x\frac{d^2f}{dx^2}$$

$$\Delta x = -\frac{\frac{df}{dx}}{\frac{d^2f}{dx^2}}(x)$$

$$\boxed{\Delta x = -\frac{f'(x)}{f''(x)}}$$

Newton Step

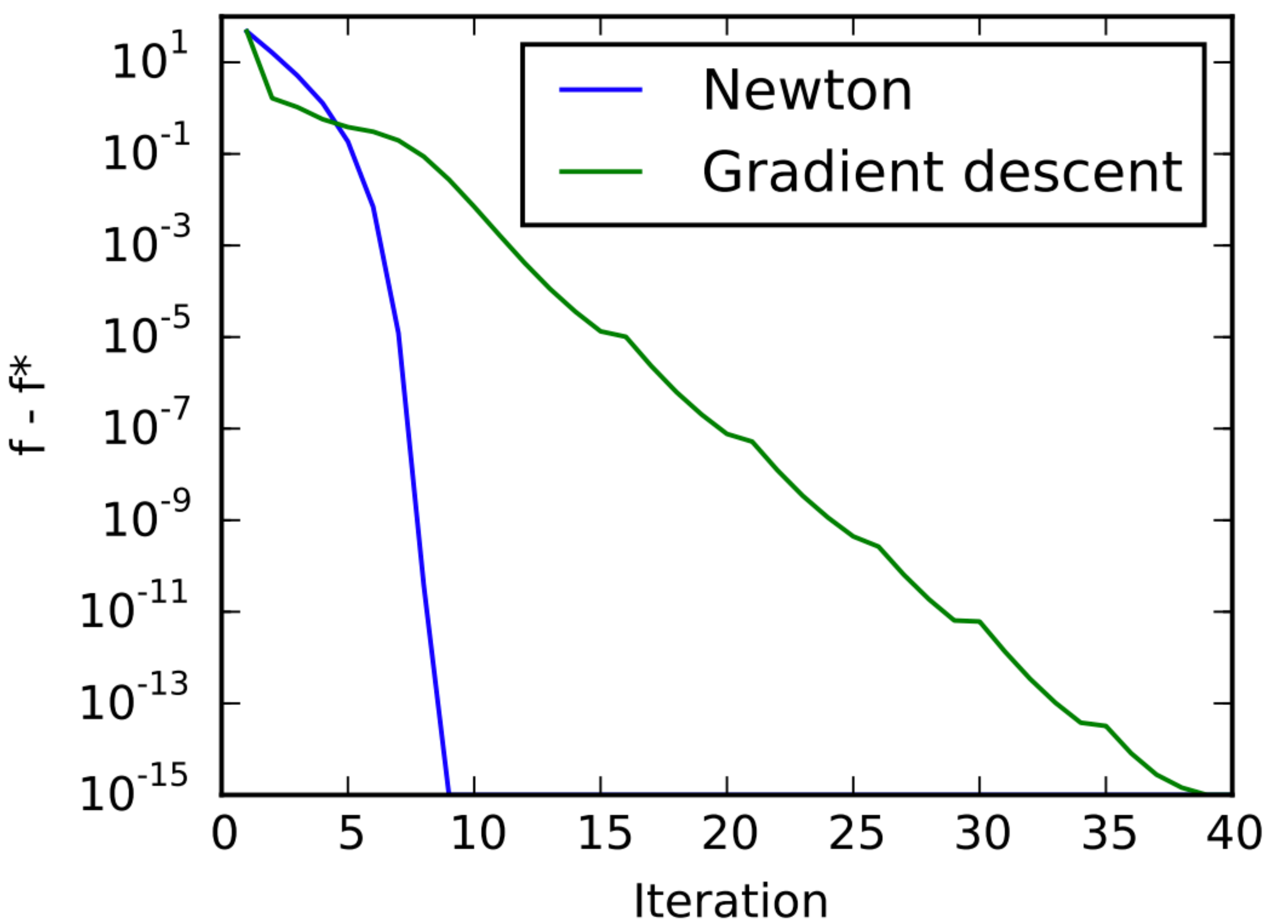


# Optimal Minimizations

## Example: Newton's method

$$f(x_1, x_2) = \exp(x_1 + 3x_2 - 0.1) + \exp(x_1 - 3x_2 - 0.1) + \exp(-x_1 - 0.1)$$

Convergence of Newton's method vs. gradient descent




# Newton step

$$\Delta x_{\mathrm{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

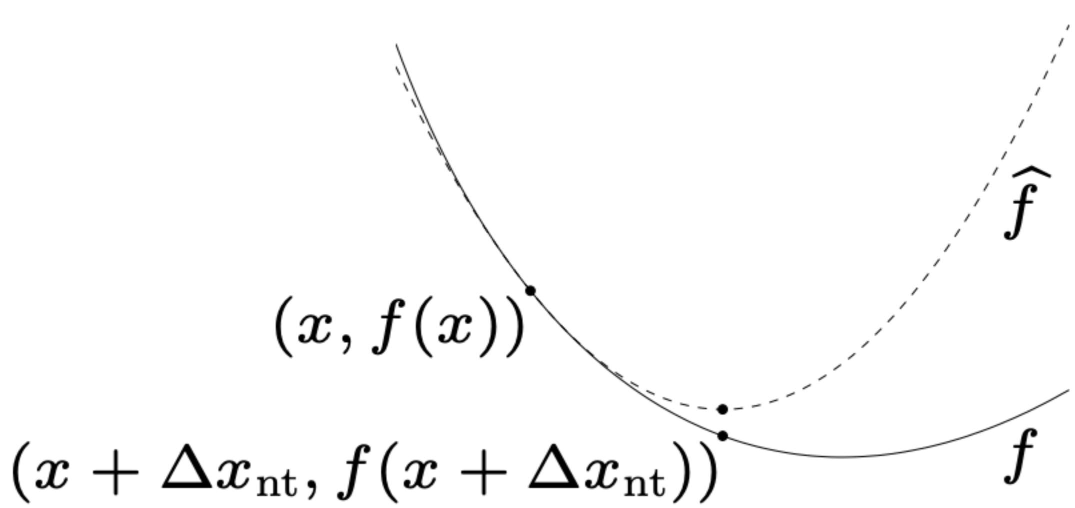
**interpretations**

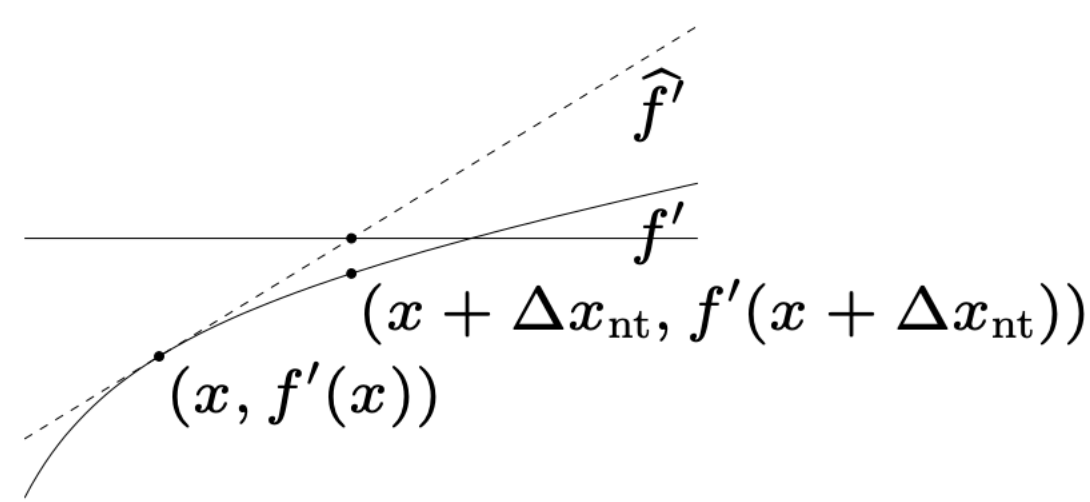
- $x + \Delta x_{\mathrm{nt}}$ minimizes second order approximation

$$\widehat{f}(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v$$

- $x + \Delta x_{\mathrm{nt}}$ solves linearized optimality condition

$$\nabla f(x+v) \approx \nabla \widehat{f}(x+v) = \nabla f(x) + \nabla^2 f(x) v = 0$$



Its also a cool dance move
that Newton did

## Newton step

$$\Delta x_{\mathrm{nt}} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

**interpretations**

- $x + \Delta x_{\mathrm{nt}}$ minimizes second order approximation

$$\widehat{f}(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2$$

- $x + \Delta x_{\mathrm{nt}}$ solves linearized optimality condition

$$\nabla f(x+v) \approx \nabla \widehat{f}(x+v) = \nabla f(x) + \nabla^2 f(x) v = 0$$