Hyper-fast gravitational wave parameter estimation

Hunter Gabbard, PhD Candidate, University of Glasgow

PyData Edinburgh Meeting

January 9th, 2019

Talk Overview

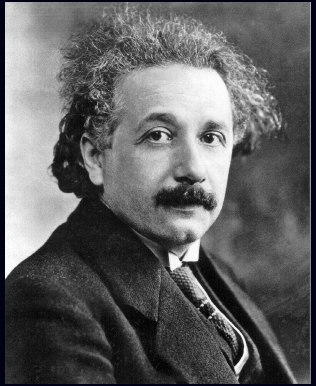
- What are gravitational waves?
- What are conditional variational autoencoders?
- What did we do?
- Where do we go from here?

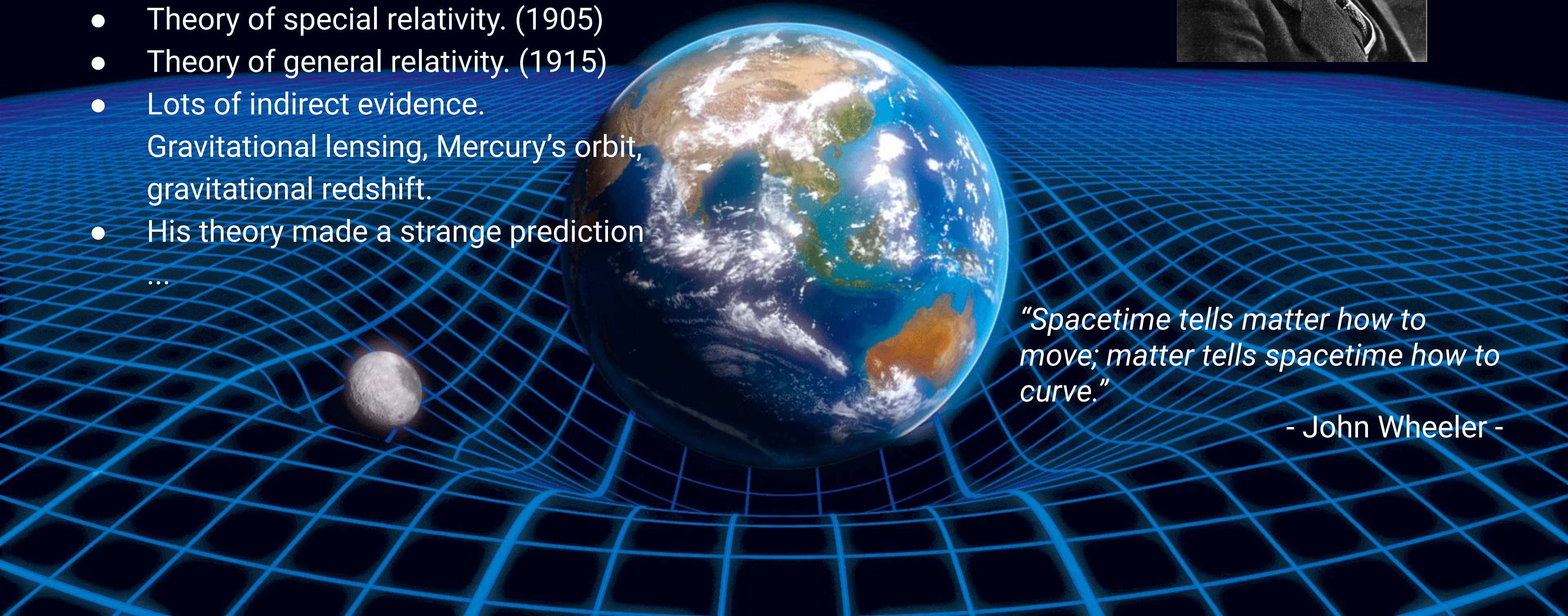
What are Gravitational Waves?

(things that go bump in the metric)

General Relativity

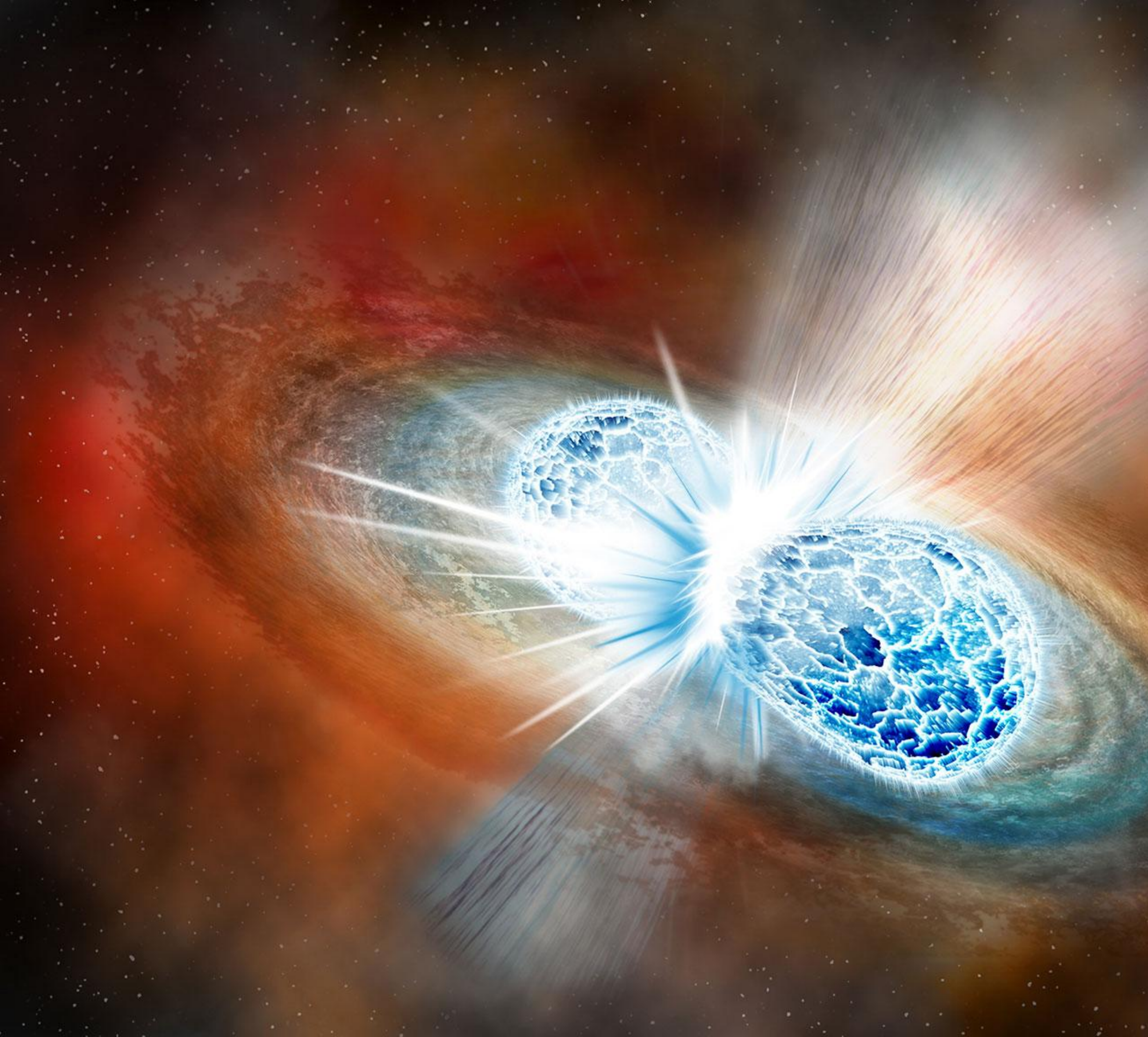
- Theory of special relativity. (1905)
- Theory of general relativity. (1915)
- Lots of indirect evidence.
  Gravitational lensing, Mercury’s orbit, gravitational redshift.
- His theory made a strange prediction
  ...





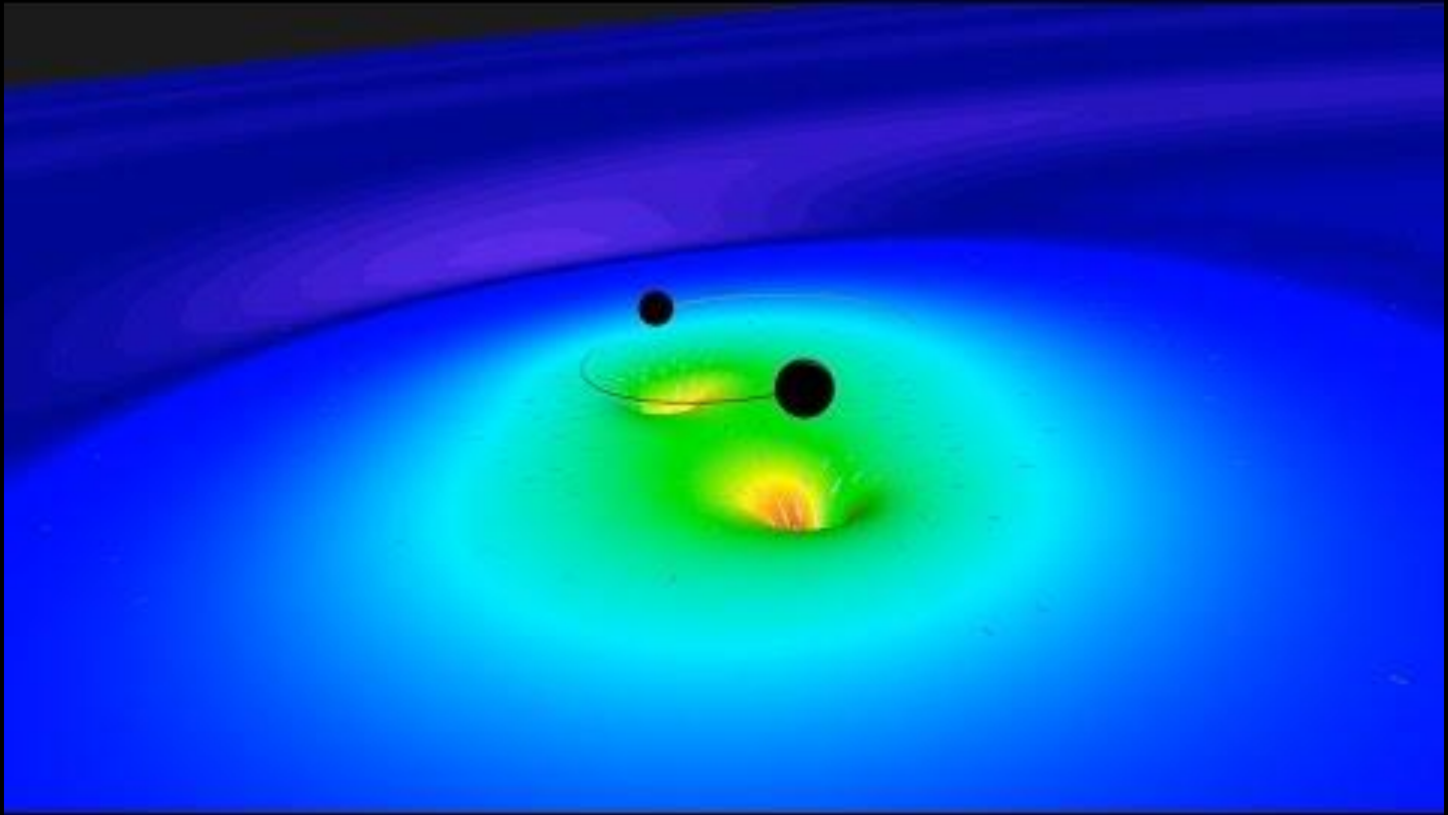
“Spacetime tells matter how to move; matter tells spacetime how to curve.”

- John Wheeler -



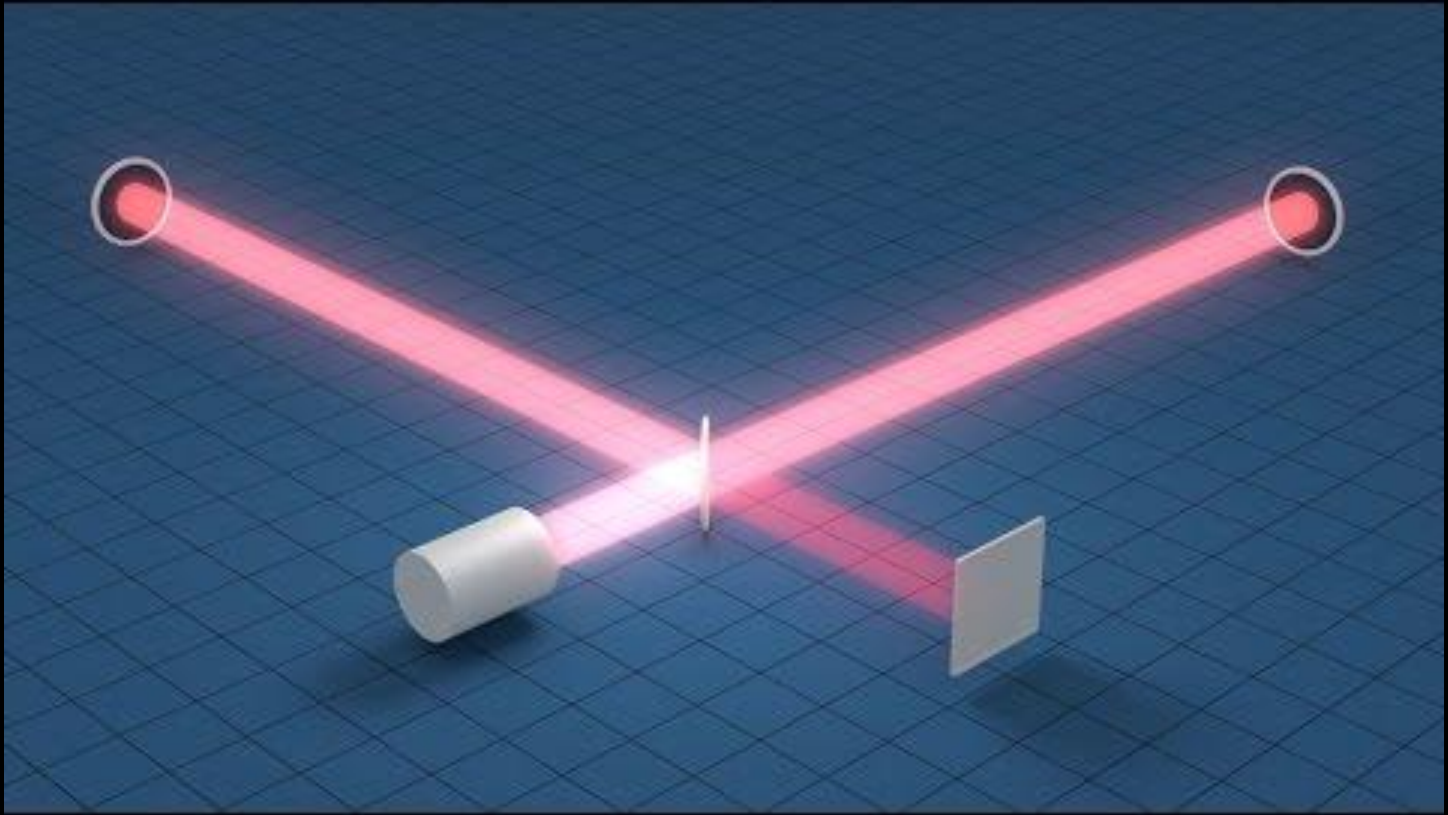
Gravitational waves

- Predicted by Einstein in 1916.
- Result from some of the most violent events in the known universe.
- Potential sources: binary black holes, binary neutron stars, supernovae, the big bang … other unknown?
- If detected, can do whole new form of astronomy without using light!



How do we detect them?

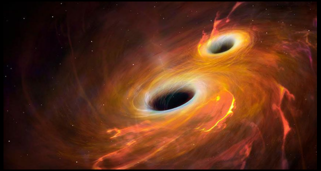
(the world’s most precise ruler)

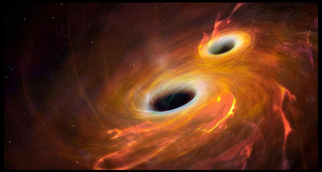


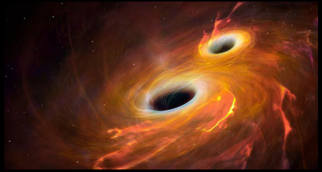

Scots scienti
played key pa
in 'monumen
the guardian
So it turns out Einstein was right all along ...
Questions Google can't answer
i
OTSMAN
THE
SCOTLAND
Country life
Black hole pro
Huge explosion in a galaxy far, far away confirms Einstein's space-time theory
eds fresh
of universe
The Oregonian
KIRKUK
Kurdish-held city now in Iraqi control
ALWAYS ON OREGONLIVE.COM Tuesday, October 17, 2017 $1.50
ASTRONOMY
Scientists strike gold
The detection of two neutron stars colliding holds promise for more discovery
Waves and radiation:
A new astronomy:
Among their discoveries:
Astronomical first:
The New York Times
Late Edition
NEW YORK, FRIDAY, FEBRUARY 12, 2016
$2.50
WITH FAINT CHIRP, SCIENTISTS PROVE EINSTEIN CORRECT
A RIPPLE IN SPACE-TIME
An Echo of Black Holes Colliding a Billion Light-Years Away
By DENNIS OVERBYE
A worker installed a baffle in 2010 to control light in the Laser Interferometer Gravitational-Wave Observatory in Hanford, Wash.
Long in Clinton's Corner, Blacks Notice Sanders
Courted Hard in South Carolina, Loyalists Listen Closely
By RICHARD FAUSSET
Last Occupier In Rural Oregon Is Coaxed Out
Reginald Abraham, left, an organizer for Hillary Clinton, at a barbershop in Orangeburg, S.C.
$2.00
THE NATION'S NEWS
FEBRUARY 12 - 14, 2016
USA TODAY WEEKEND
GRAVITATIONAL WAVES CONFIRMED
'A WHOLE NEW WINDOW ON THE UNIVERSE'
NEWSLINE
IN NEWS
Clinton, Sanders agree systems in U.S. are broken
In Japan, the girls give guys gifts on Valentine's Day
IN MONEY
Fed's Yellen says negative rates not 'off the table'
IN SPORTS
Gordon shifts into his new TV career
IN LIFE
Jagger, Scorsese offer look at sex, drugs, rock 'n' roll
Discovery affirms Einstein theory
Traci Watson
RUSSIA, U.S. REACH DEAL IN SYRIA WAR
Cease-fire could begin in a week, but it's far from certain
Jim Michaels
White House bans guns — finally
Sign is posted outside guard shack 28 years after law enacted
Gregory Korte
The "No firearms" sign now stands at the White House.
itational waves detected again
Keine Neonazis in der DDR? Vor 30 Jahren wurden sie unüberhörbar – Seite 3
DER TAGESSPIEGEL
Kosmische Kollision
Wahlen befeuern Debatte über Kurs der Union
The Boston Globe
TUESDAY, OCTOBER 17, 2017
When stars collide
Tax overhaul bill uncertain for '17
Vertex pledges $500m over 10 years for charity
Receive and consent
A COSMIC BLAST THAT SHAKES SCIENCE CIRCLES
THE NEW INDIAN EXPRESS
Guns blazing, Modi lobs progress dare at Cong
After fact check, Modi figure powered down by a crore
Air India
More Pride... More Value

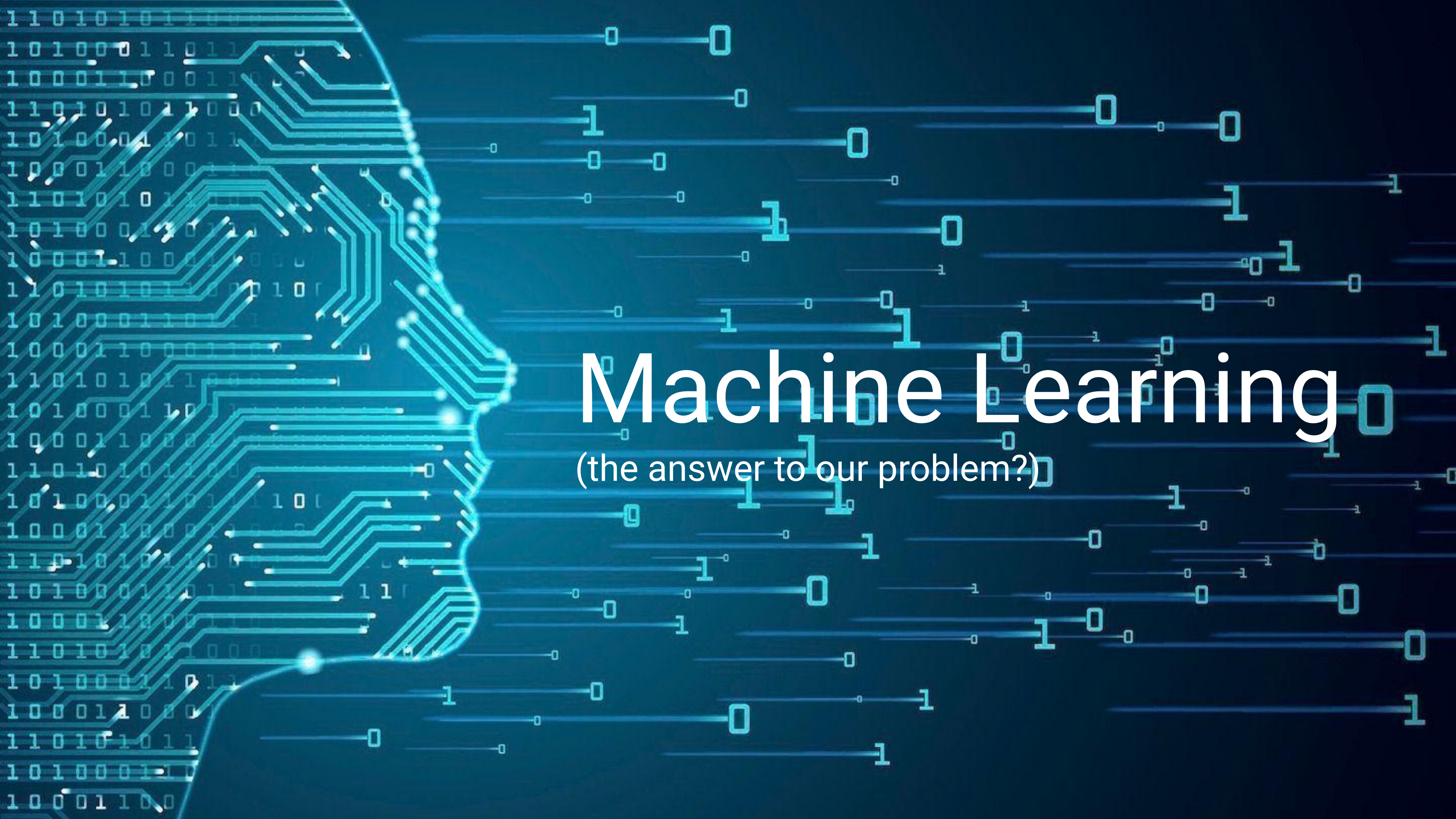

ALFR. NOBEL
NAT. MDCCC XXXIII
OB. MDCCC XCVI

So what now? ...







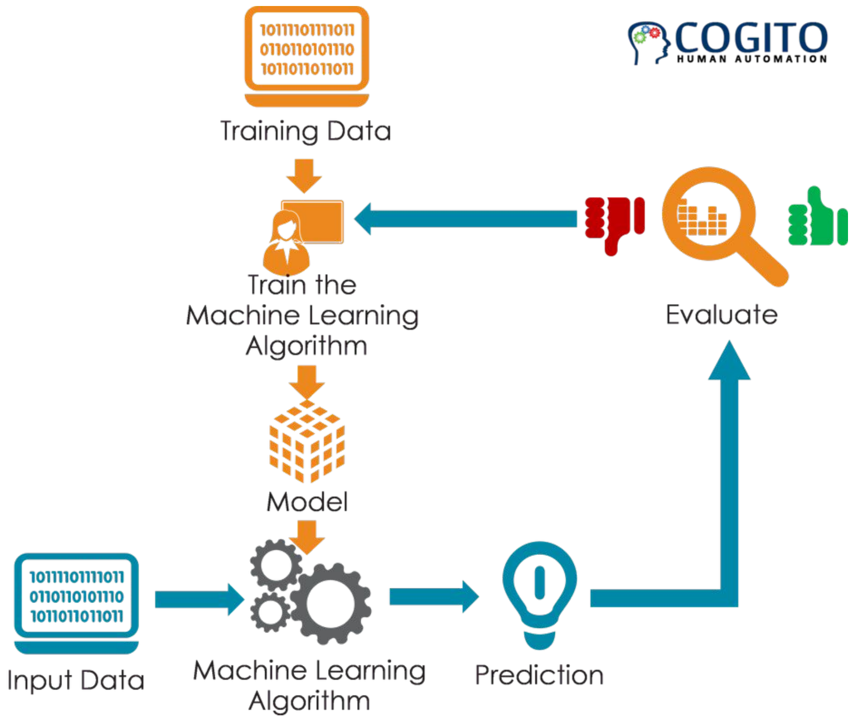


Machine Learning

(the answer to our problem?)

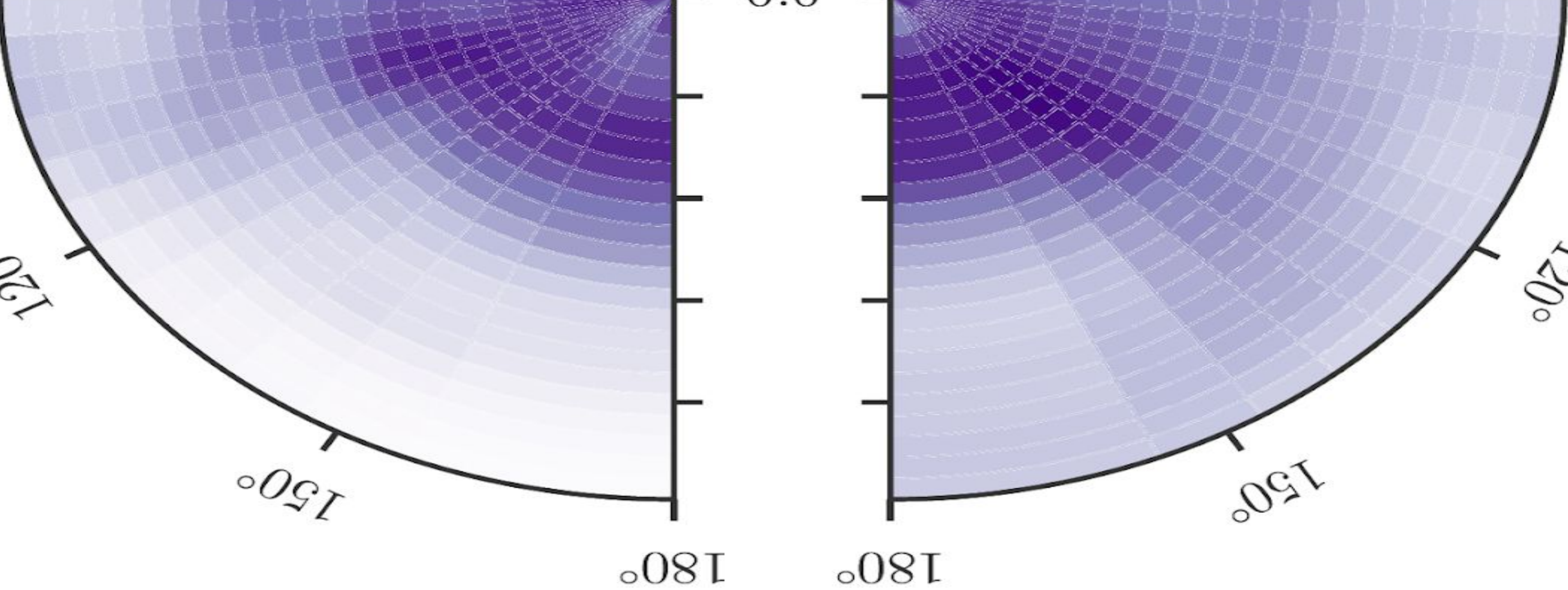
What is Machine Learning?

- Good question Hunter!
- Is it just a useful black box? NO!
- It is a complex system composed of simple operations that is initially “stupid”. Through the initial choice of incentive the system can be trained to do certain tasks.
- Often referred to as a simple function approximator.



Why use machine learning?

- Speed
    - Front loading the cost - leads to rapid real-time analysis.
    - Simply saves CPU cycles and therefore cold hard cash.
    - Allows us to keep up with the detections.
- It’s just better (maybe)
    - Existing techniques are (near) optimal when you know your model (signal and noise).
    - ML has the potential to learn what we have approximated or incorrectly assumed.
    - If faster than computationally limited searches - can be made more sensitive.



Parameter estimation

(what exactly did we just detect?)

Bayesian Inference

- We rely heavily on Bayesian parameter estimation.

$$\underbrace{p(\boldsymbol{\theta}|\boldsymbol{x},I)}_{\text{posterior}} = \frac{\overbrace{p(\boldsymbol{x}|\boldsymbol{\theta},I)}^{\text{likelihood}}\,\overbrace{p(\boldsymbol{\theta}|I)}^{\text{prior}}}{\underbrace{p(\boldsymbol{x}|I)}_{\text{evidence}}}$$

- We must make (motivated) assumptions regarding our noise distribution and the prior values of of the parameters.

$$\log p(\boldsymbol{x}|\boldsymbol{\theta},I) = \exp\left[-\sum_{j=0}^{N-1}\frac{|\tilde{x}_j-\tilde{h}_j(\boldsymbol{\theta})|^2}{TS_h(f_j)}\right] + \text{const}$$

- In practice, obtaining posterior probability densities on system parameters is done via Monte-Carlo techniques (MCMC, nested sampling).

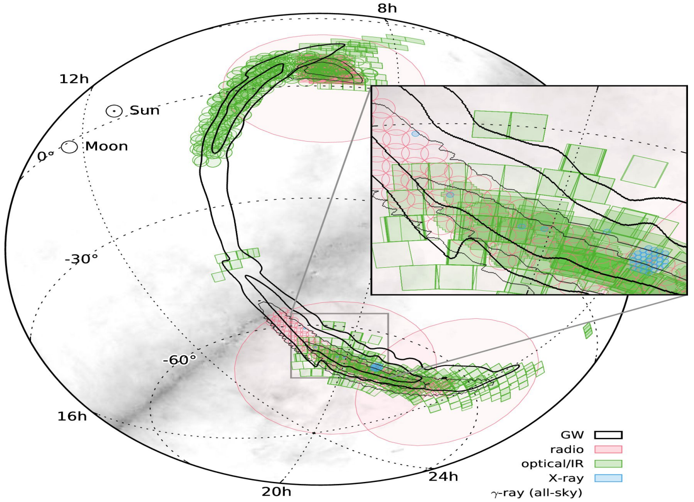
https://lscsoft.docs.ligo.org/bilby/

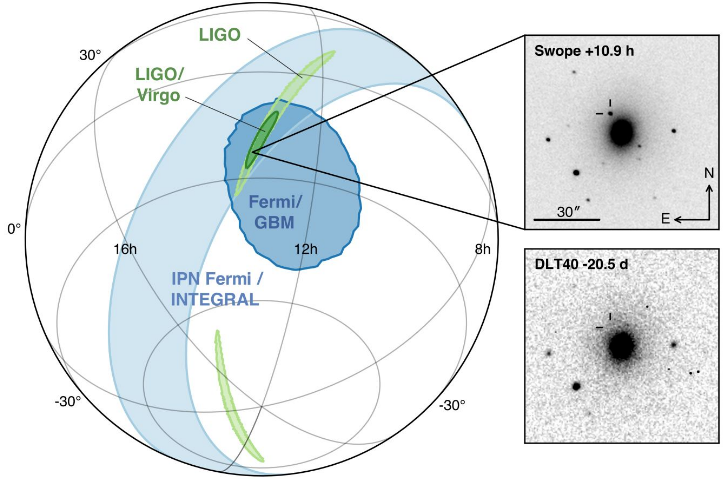
L. Singer & L. Price, PRD 93, 2, 024013 (2016)

Motivation

- Existing Bayesian parameter estimation is optimal but very slow
- For transient GW events and multi-messenger astronomy it is crucial that we produce data products very quickly.

LIGO-Virgo Collaboration, ApJ 848, 2, L12, 59 (2017)





LIGO-Virgo Collaboration, ApJ Volume 826, 1, L13, 8 (2016)

Conditional Variational Autoencoders

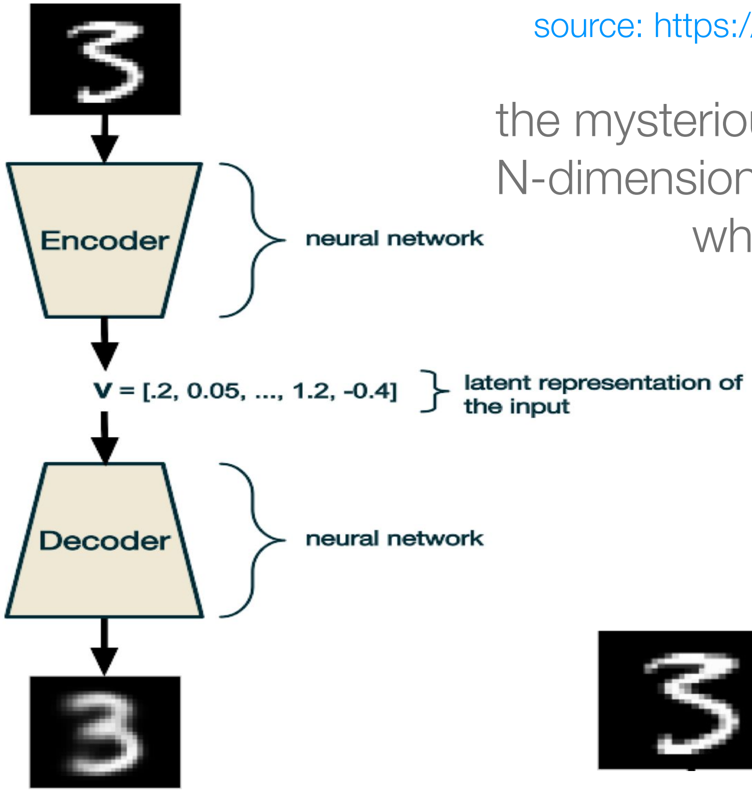
(it’s not magic, I promise)

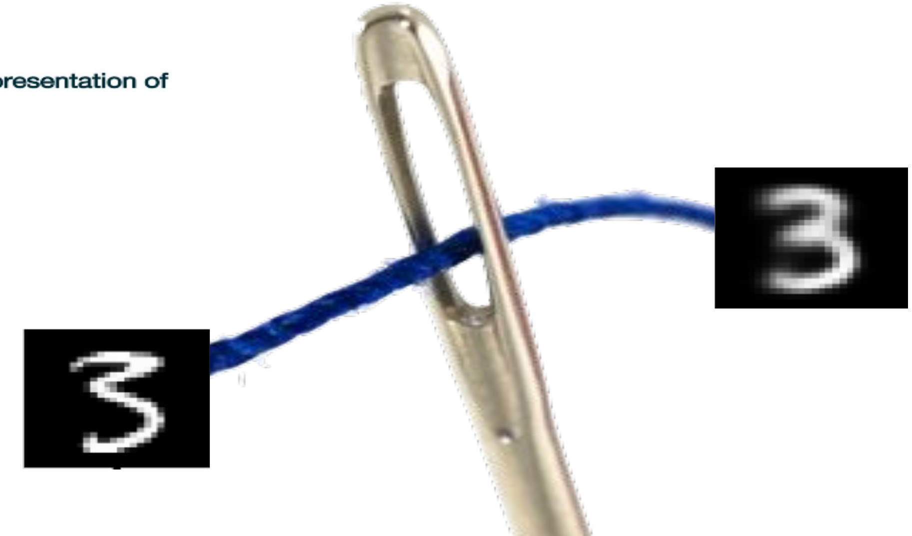
First - an autoencoder

source: https://ijdykeman.github.io/ml/2016/12/21/cvae.html

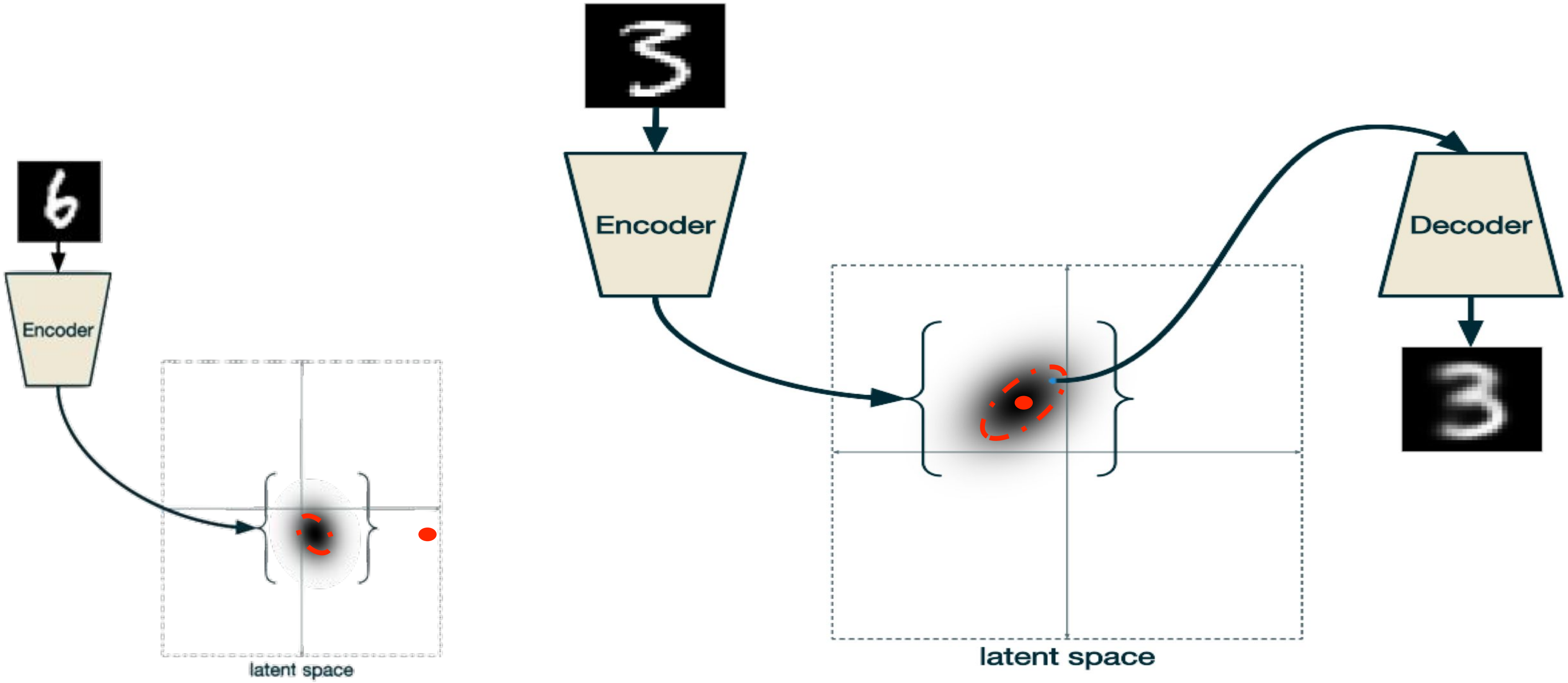
the mysterious “latent” space is jargon for some N-dimensional non-physical parameter space in which to represent your data

loss based on input/output similarity





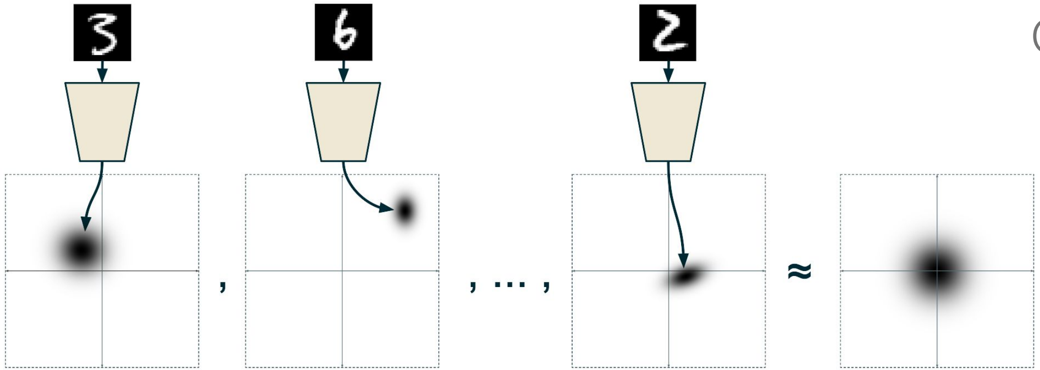
Next - a variational autoencoder

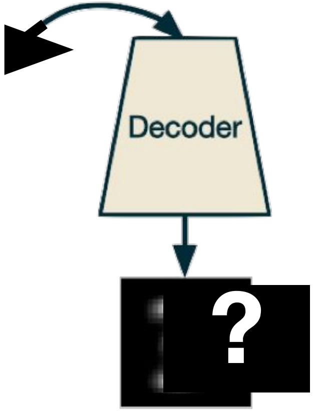


source: https://ijdykeman.github.io/ml/2016/12/21/cvae.html

Next - a variational autoencoder

The loss function incorporates a KL-divergence term testing the Gaussianity of the total distribution in the latent space





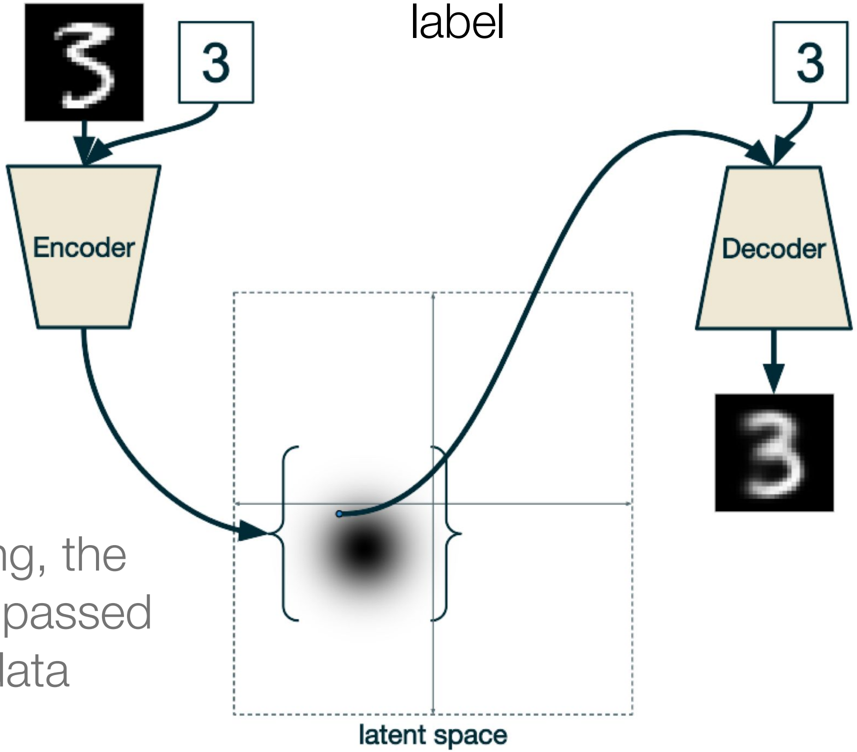
random draw
can’t produce particular numbers on command

source: https://ijdykeman.github.io/ml/2016/12/21/cvae.html

Then - a conditional variational autoencoder

label

label



Within training, the true “label” is passed with the data

now you can ask for a random “7” or “4”, etc…

source: https://ijdykeman.github.io/ml/2016/12/21/cvae.html

Some of the components of our scheme

The GW data y

we generate lots of these

The GW parameters x

from this

The prior $p(x)$

we never have to evaluate this

The likelihood $p(y|x)$

we want this

The latent space parameters z

The posterior $p(x|y)$

we are going to use 3 neural networks to model each of these 3 functions

The recognition function $q_\phi(z|x,y)$

The target distribution $p_\theta(x|y) = \int dz\, r_{\theta_1}(z|y) r_{\theta_2}(x|z,y)$

Gabbard *et al*, arXiv:1909.06296 (2019)

The scheme

Loss function:

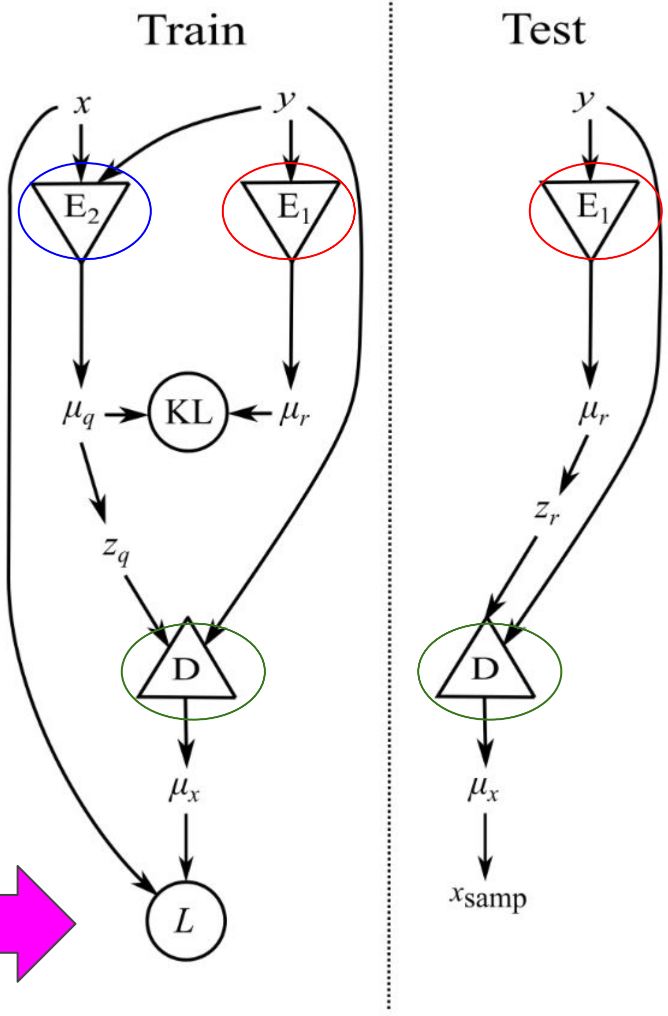
- Start of derivation

$$H(p,r) = -\int dx\, p(x|y)\log r_\theta(x|y)$$

... Here be math dragons ...

- End of derivation

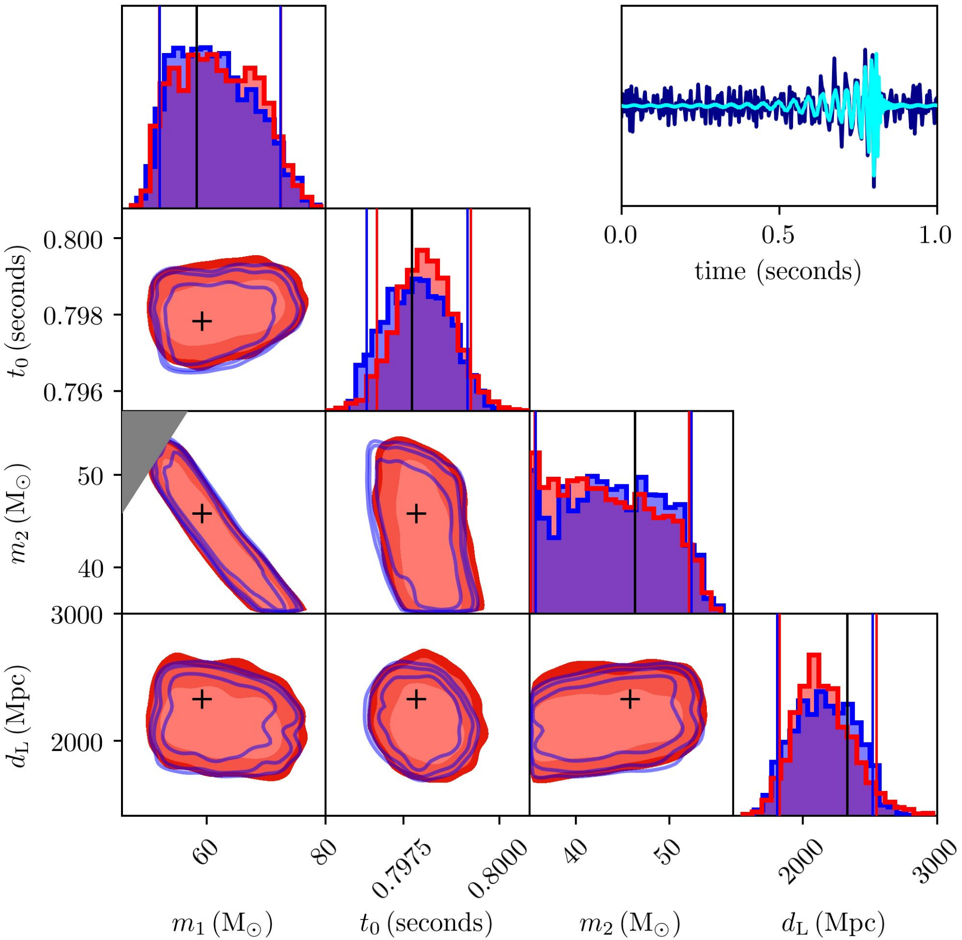
$$H \lesssim -\frac{1}{N}\sum_{n=1}^{N}\left[\log r_{\theta_2}(x_n|z_n,y_n)\right. \\ \left. -\mathrm{KL}\left[q_\phi(z|x_n,y_n)\|r_{\theta_1}(z|y_n)\right]\right].$$



What did we find out?

(it’s fast)

Posterior comparisons



Red: Our method
Blue: Other method

The speed

TABLE I. Durations required to produce samples from each of the different posterior sampling approaches.

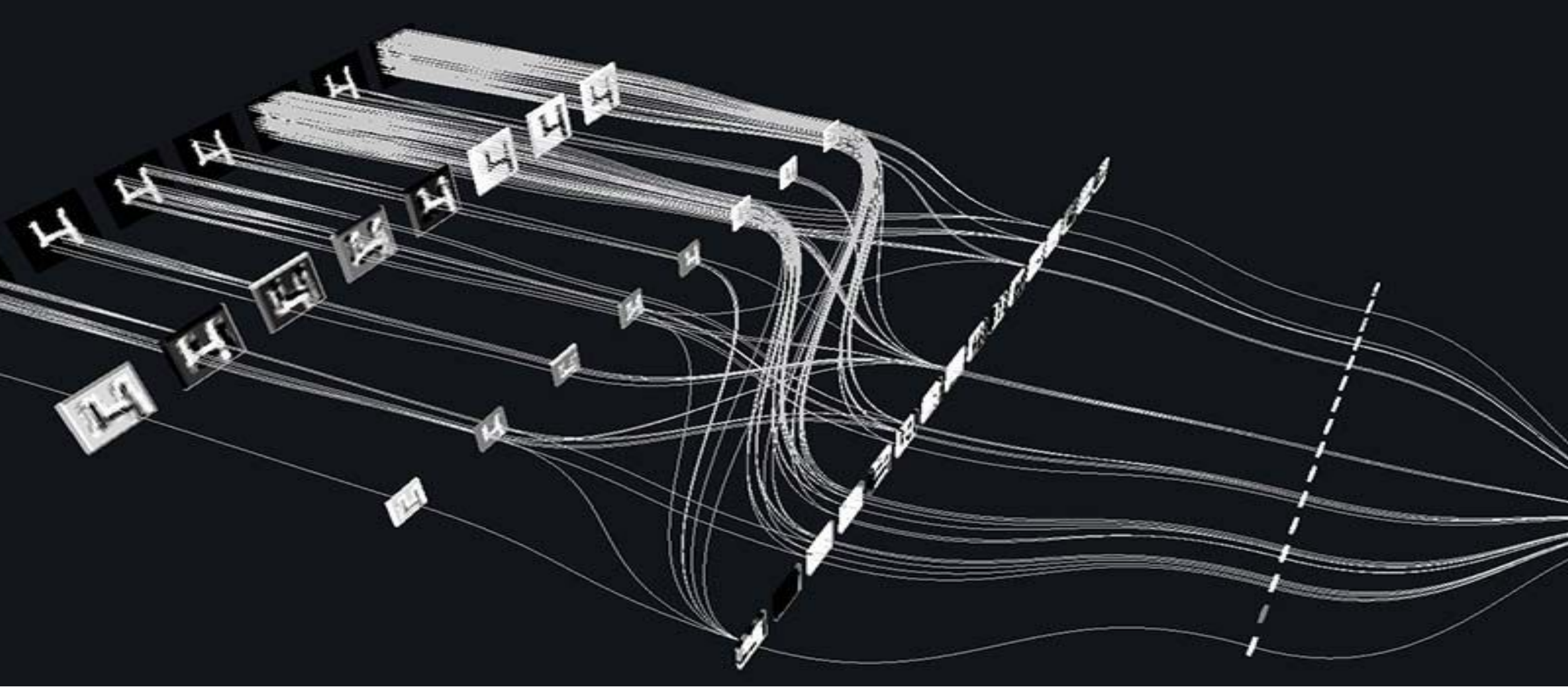
sampler	run time (seconds) min	max	median	ratio $\frac{\tau_{\text{VItamin}}}{\tau_X}$
Dynesty[a]	602	1538	774[b]	2.6×10^{-6}
Emcee	2005	11927	4351	4.6×10^{-7}
Ptemcee	3354	12771	4982	4.0×10^{-7}
Cpnest	1431	5405	2287	8.8×10^{-7}
`VItamin`[c]		$\mathbf{2 \times 10^{-3}}$		1

[a] The benchmark samplers all produced $\mathcal{O}(3000 - 10000)$ samples dependent on the default sampling parameters used.

[b] The reader may note that benchmark sampler run times are a few orders of magnitude lower than what is typical of a complete BBH analysis ($\mathcal{O}(10^5 - 10^6)$ seconds). This is primarily due our use of a reduced parameter space, low sampling rate and choice of sampler hyperparameters.

[c] For the `VItamin` sampler 3000 samples are produced as representative of a typical posterior. The run time is independent of the signal content in the data and is therefore constant for all test cases.

Gabbard *et al*, arXiv:1909.06296 (2019)



Conclusions

The take home message

Variational Inference - Future work

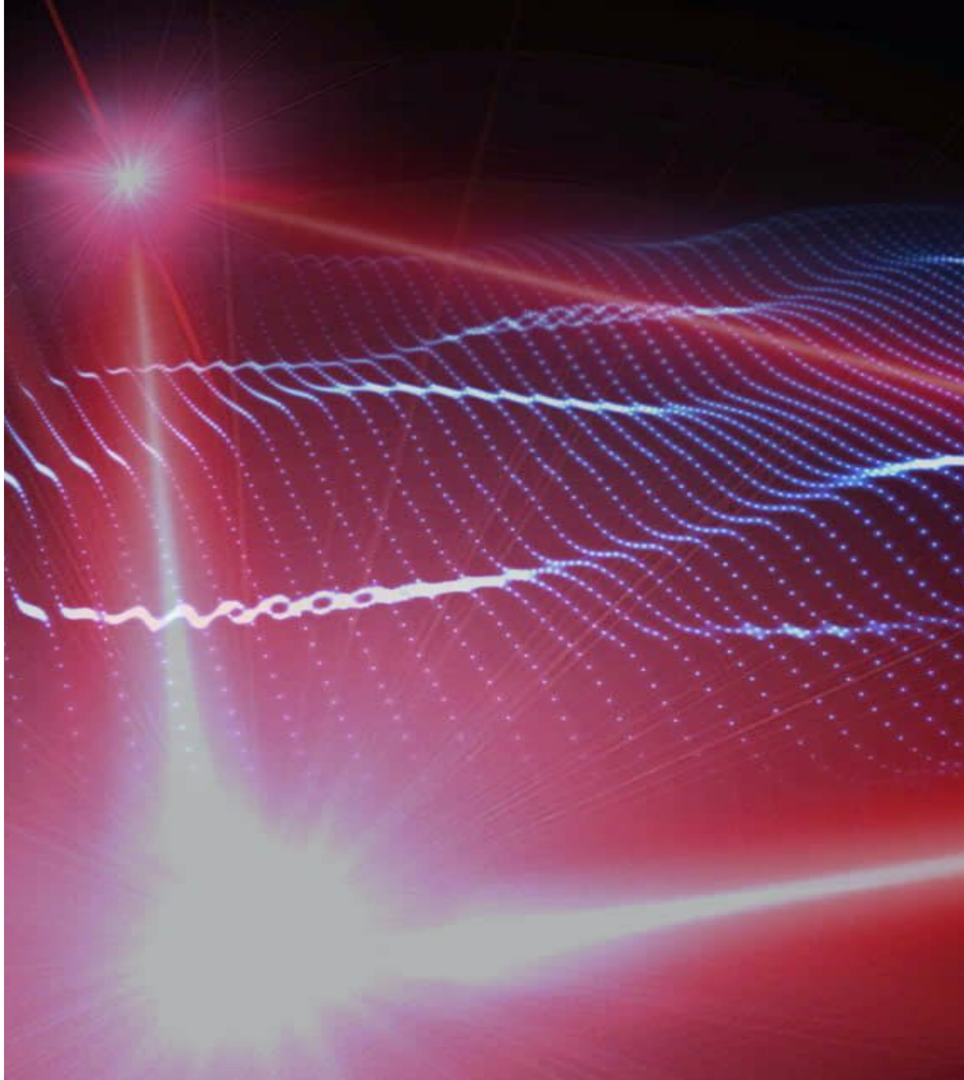
- We've shown (and so have another group) [Chua & Valisneri, arXiv:1909.05966 (2019)] that variational inference is a powerful tool.

- Extending this to more realistic cases is the next step.

- The ultimate aim is to have this working on CBC signals containing matter.

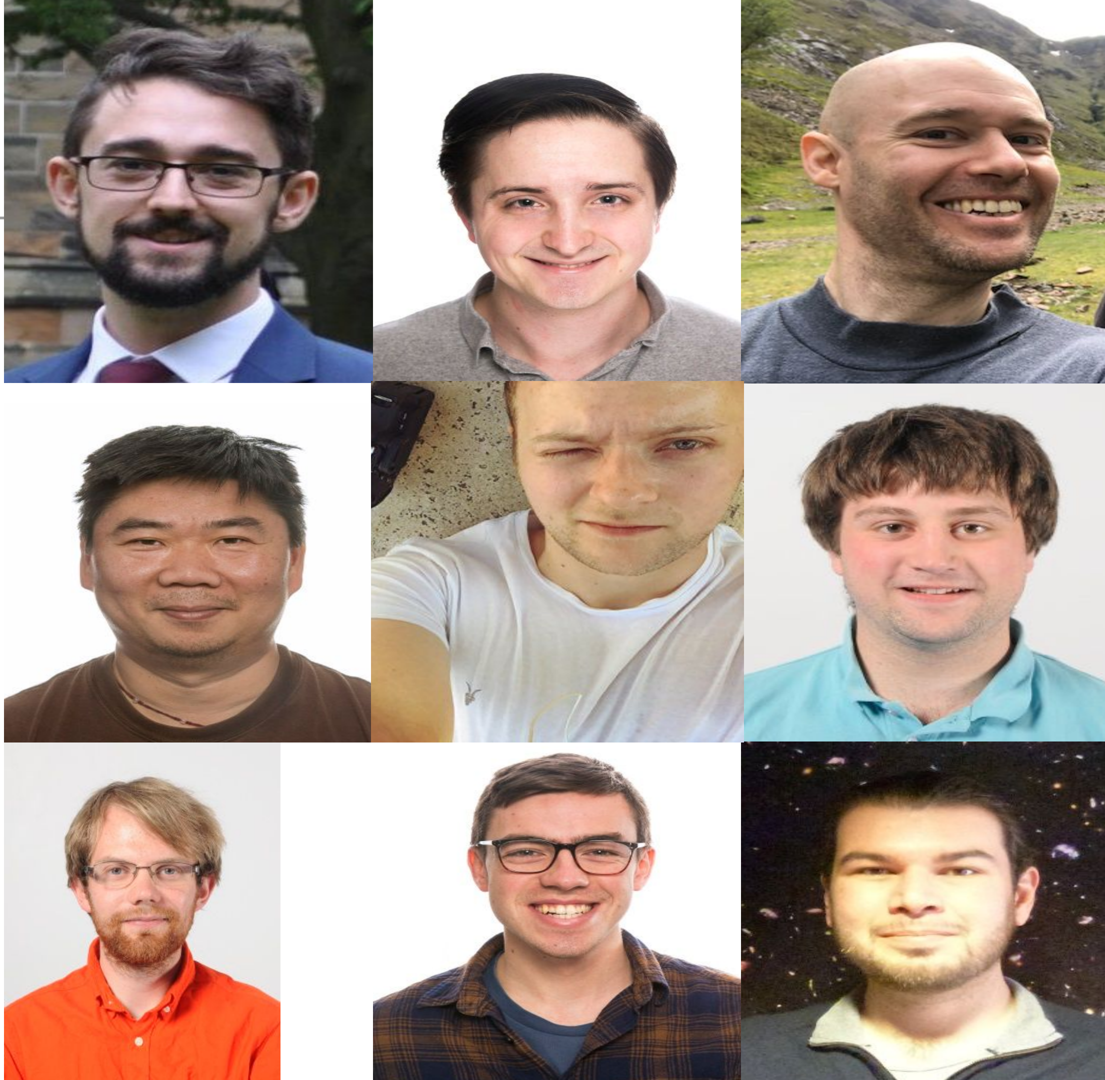
- Our pipeline is called VItamin and is available to play with here

https://github.com/hagabbar/VItamin



Summary

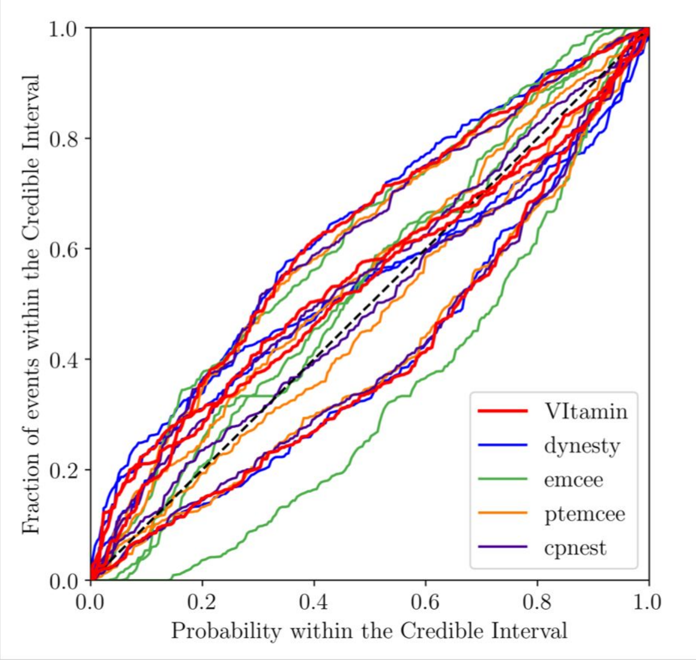
- We covered the basics of gravitational wave astronomy
- We then covered some ML basics
- We covered variational autoencoders
- We finished off with variational inference for Bayesian parameter estimation.



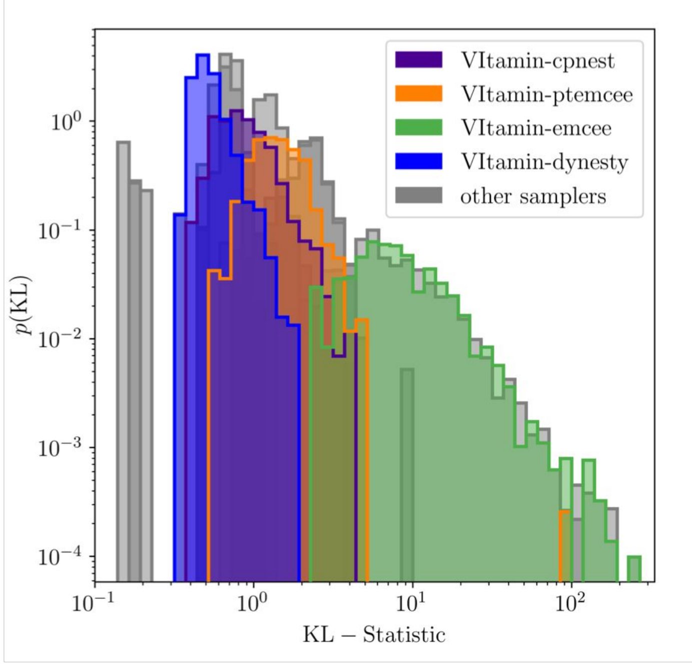
Thank you for your attention!

Validation

p-p plot



KL-divergence distributions



Gabbard ***et al***, arXiv:1909.06296 (2019)