Facing the challenges of modern cosmological surveys with deep learning

François Lanusse

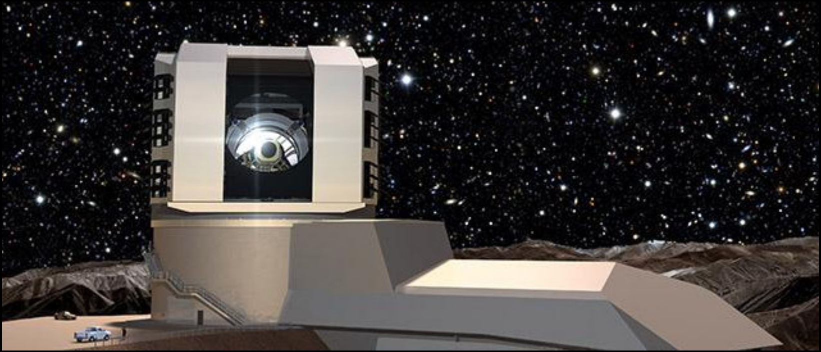
January 16, 2017

McWilliams Center for Cosmology
Carnegie Mellon University

Carnegie Mellon University
McWilliams Center for Cosmology

LSST DESC
Dark Energy Science Collaboration

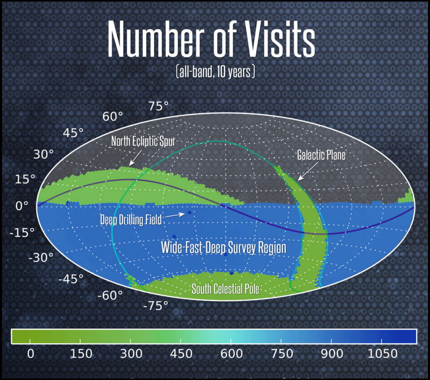
the Large Synoptic Survey Telescope



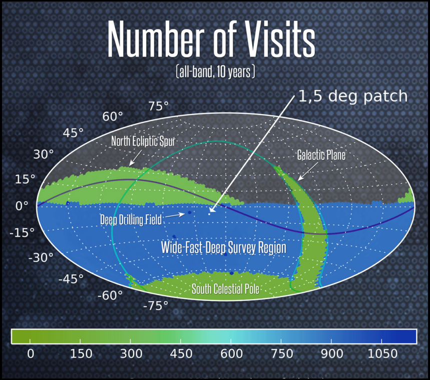
LSST in a few numbers

- 1000 images each night, each one is 3.2 GB and 40 full moons $\Longrightarrow$ 15 TB/night for 10 years
- Covers 18,000 square degrees (40% of the sky)
- Tens of billions of objects, each one observed $\sim$ 1000 times

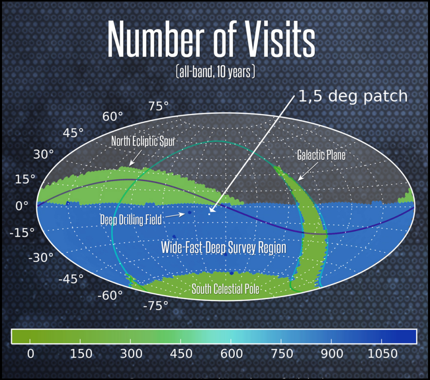
what does it look like ?



what does it look like ?

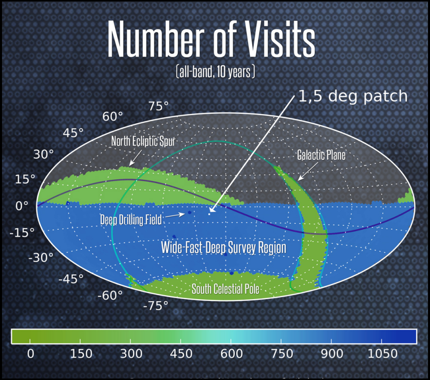


what does it look like ?



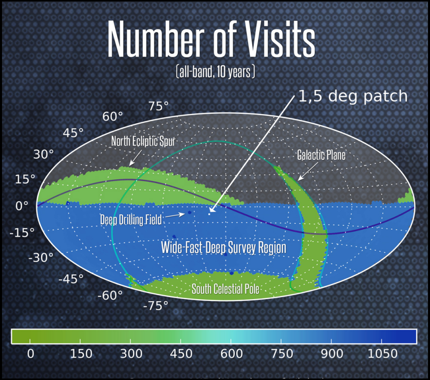
HSC-SSP Data Release 1

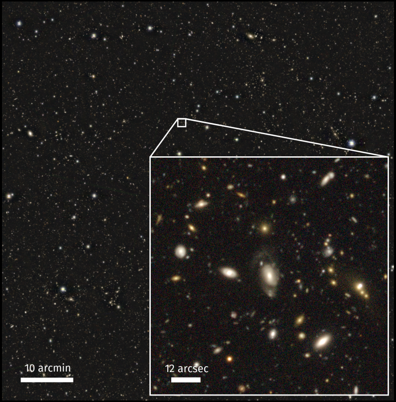
what does it look like ?



HSC-SSP Data Release 1

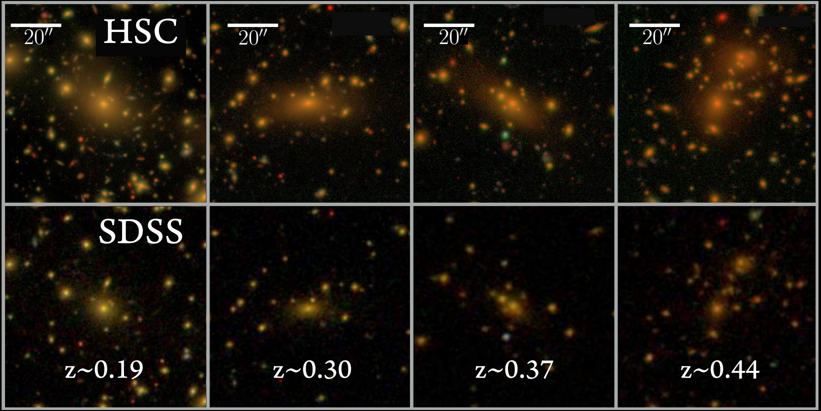
what does it look like ?





HSC-SSP Data Release 1

what does it look like ?



Huang et al. (2017), arXiv:1707.01904

the challenge for modern surveys

$\Longrightarrow$ Modern surveys will provide large volumes of high quality data

the challenge for modern surveys

$\Longrightarrow$ Modern surveys will provide large volumes of high quality data

A Blessing

- Unprecedented statistical power
- Great potential for new discoveries

the challenge for modern surveys

$\Longrightarrow$ Modern surveys will provide large volumes of high quality data

A Blessing

- Unprecedented statistical power
- Great potential for new discoveries

A Curse

- Existing methods are reaching their limits (computational cost, accuracy) at every step of the science analysis
- Control of systematic uncertainties becomes paramount

the challenge for modern surveys

$\Longrightarrow$ Modern surveys will provide large volumes of high quality data

A Blessing

- Unprecedented statistical power
- Great potential for new discoveries

A Curse

- Existing methods are reaching their limits (computational cost, accuracy) at every step of the science analysis
- Control of systematic uncertainties becomes paramount

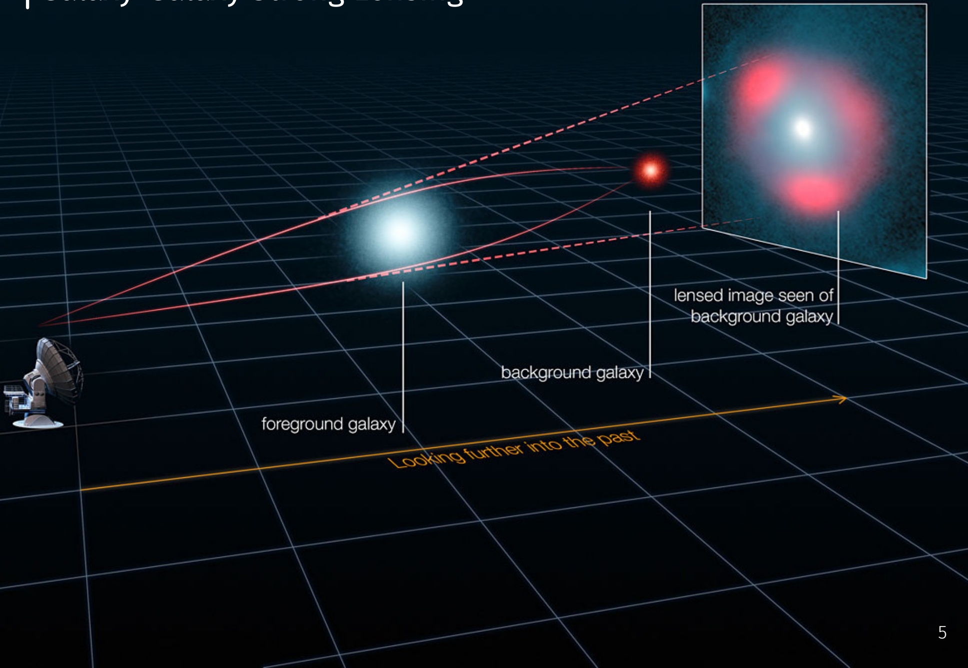
$\Longrightarrow$ Dire need for **novel data analysis techniques** to fully realize the potential of modern surveys.

Outline of this talk

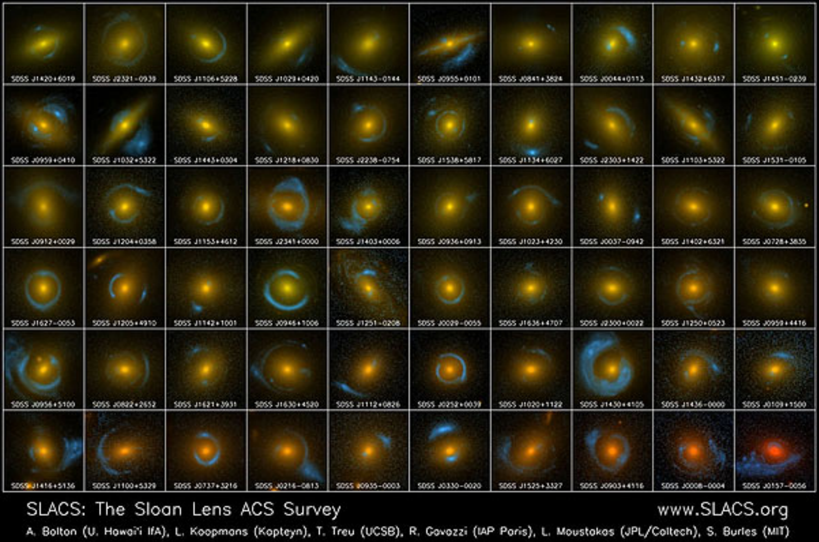
1. Deep residual networks for the detection of gravitational lenses
2. Deep Generative Models for modeling galaxy morphologies
3. Graph convolutional networks for cosmological simulations
4. Conclusion

Deep residual networks for the detection of gravitational lenses

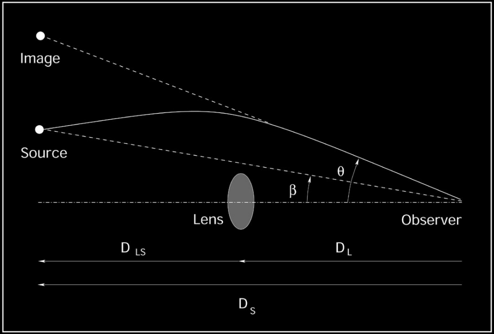
Galaxy-Galaxy Strong Lensing



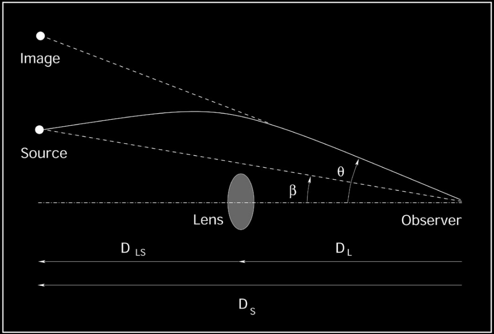
examples of strong lenses



example of application: gravitational time delays

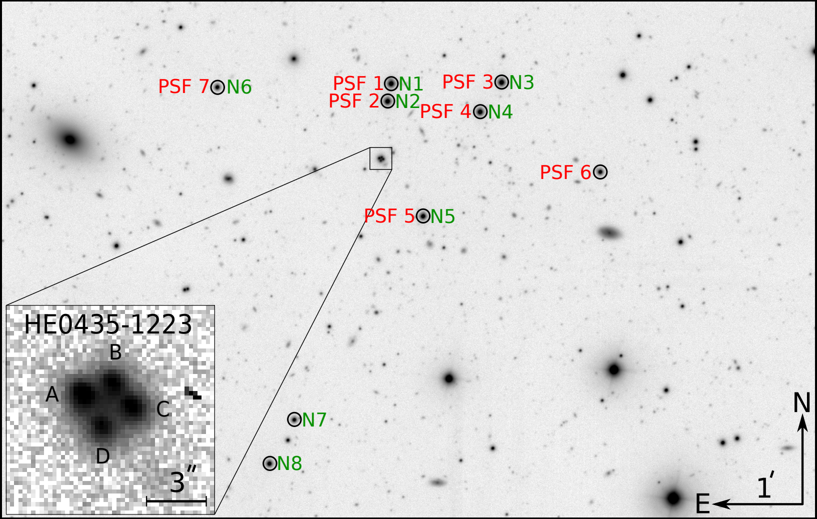


example of application: gravitational time delays

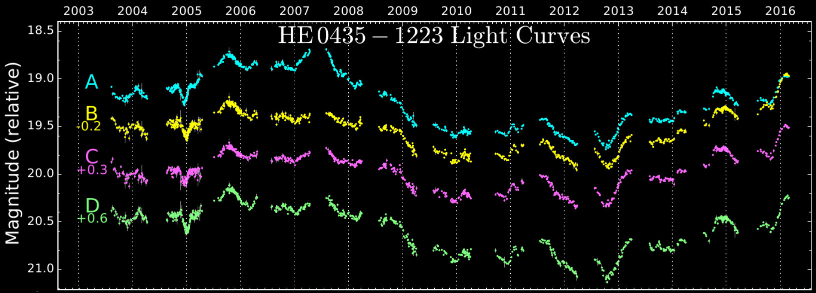


$$\Delta t_{ij} = \frac{1+z_L}{c} \underbrace{\frac{D_L\ D_S}{D_{LS}}}_{\propto H_0^{-1}} \left[\frac{(\boldsymbol{\theta}_i - \boldsymbol{\beta})^2}{2} - \psi(\boldsymbol{\theta}_i) + \frac{(\boldsymbol{\theta}_j - \boldsymbol{\beta})^2}{2} + \psi(\boldsymbol{\theta}_j) \right]$$

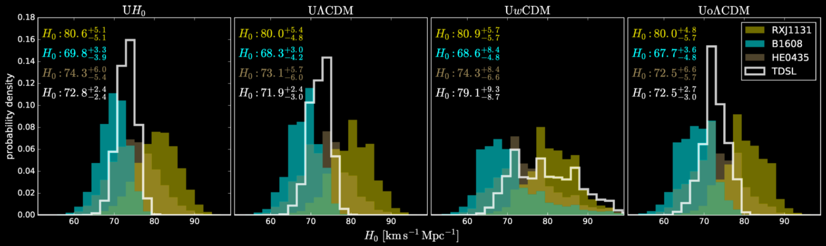
time delays of HE0435-1223 (Bonvin et al. 2017)



time delays of HE0435-1223 (Bonvin et al. 2017)

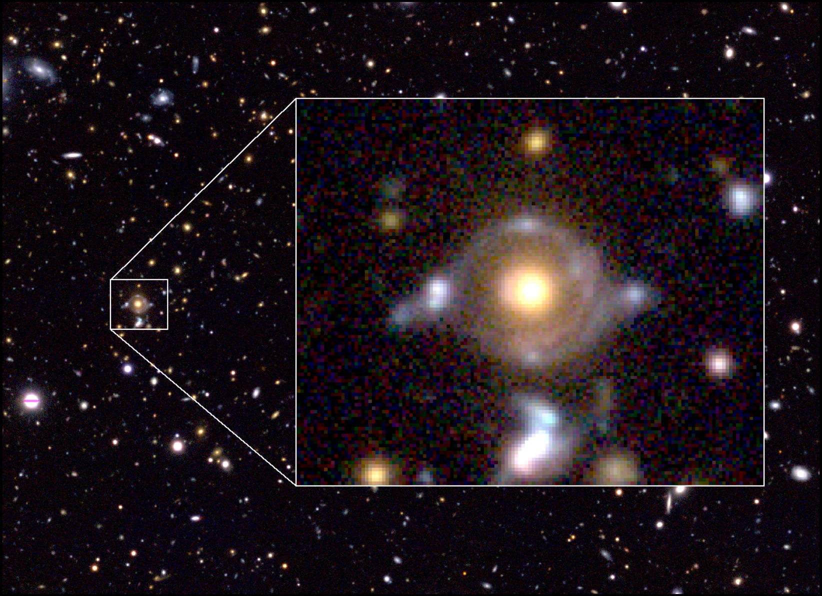


time delays of HE0435-1223 (Bonvin et al. 2017)

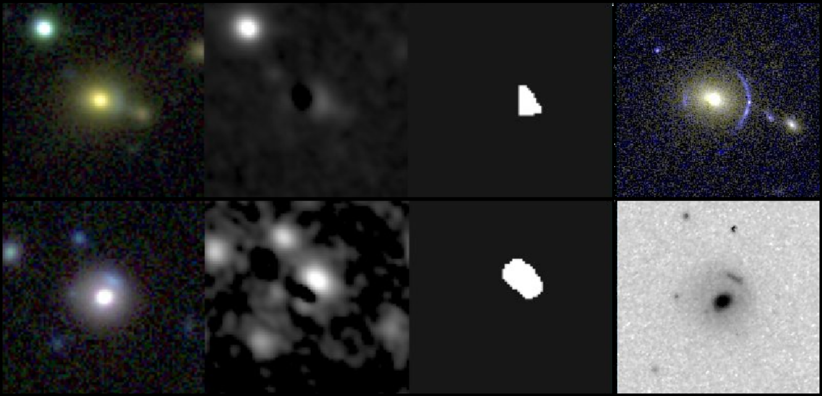


the problem: finding strong lenses

the problem: finding strong lenses

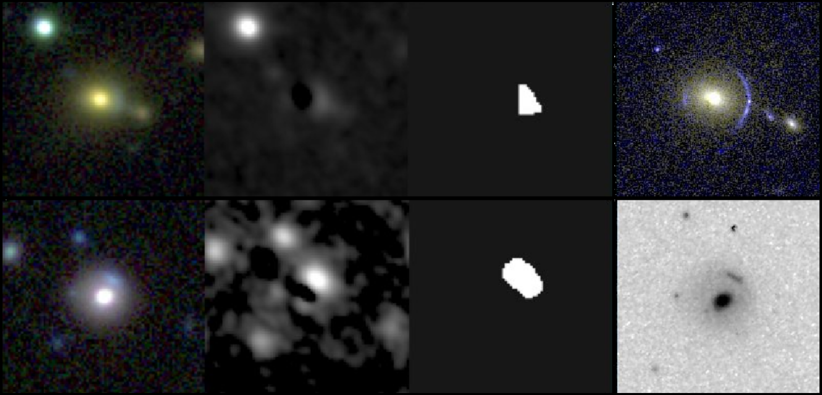


automated lens searches: RingFinder (Gavazzi et al. 2014)



gri composite $\quad g - \alpha i \quad$ detected areas $\quad$ HST images

automated lens searches: RingFinder (Gavazzi et al. 2014)

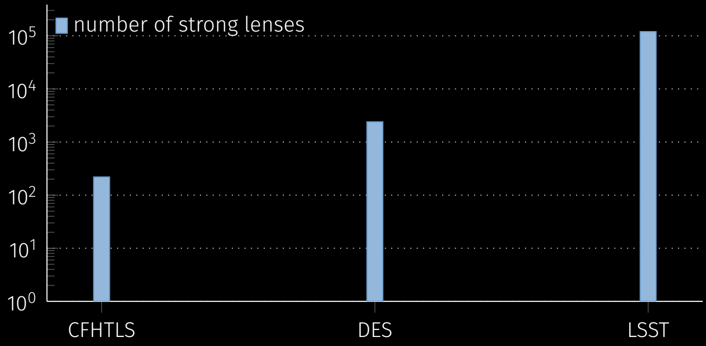


gri composite | $g - \alpha i$ | detected areas | HST images

Visual inspection time required

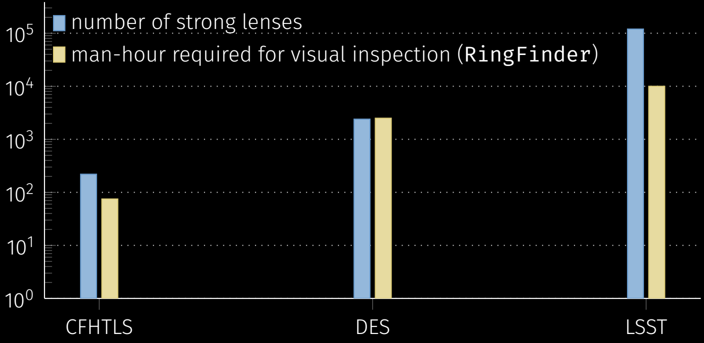
$\sim$ 30 person-minutes / deg^2

extrapolation to future surveys



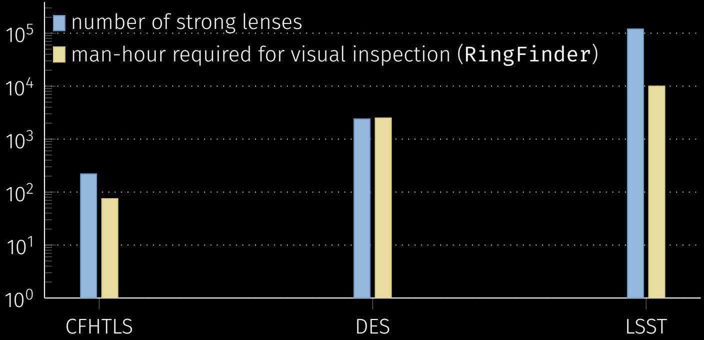
Gavazzi et al. (2014), Collett (2015)

extrapolation to future surveys



Gavazzi et al. (2014), Collett (2015)

extrapolation to future surveys

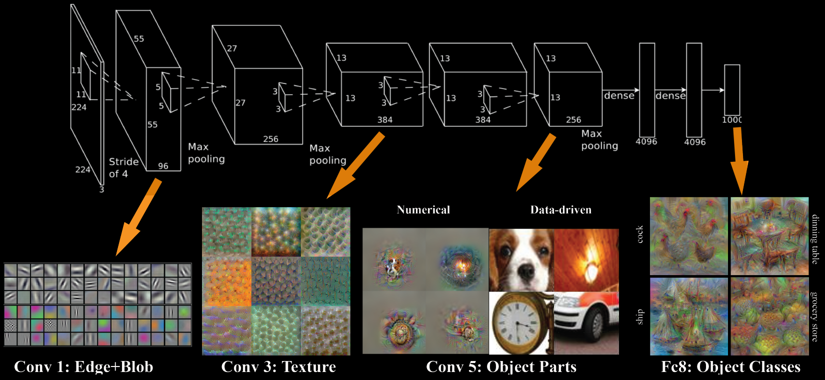


Gavazzi et al. (2014), Collett (2015)

$\Longrightarrow$ LSST would require an estimated 10^4 man-hours.

How can we robustly detect these rare objects without needing an army of grad students ?

Conventional Convolutional Neural Network



residual learning

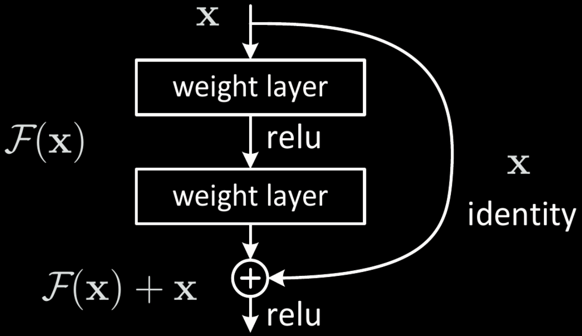


Image credit: He et al. (2015)

- Learning the difference to the identity (He et al. 2015)

residual learning

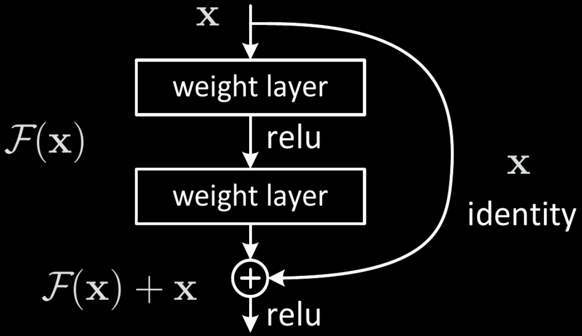
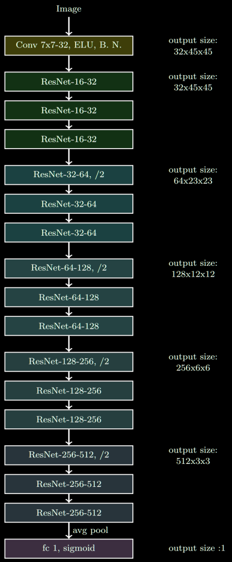


Image credit: He et al. (2015)

- Learning the difference to the identity (He et al. 2015)
- Easier to initialize and to train in deep architectures ($>$ 1000 layers)

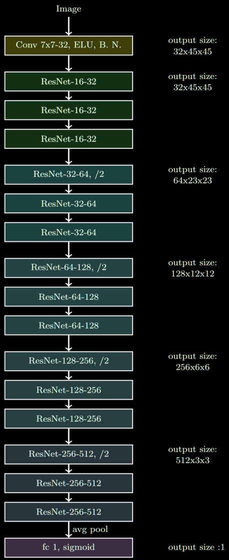
CMU DeepLens: deep residual learning for strong lens finding



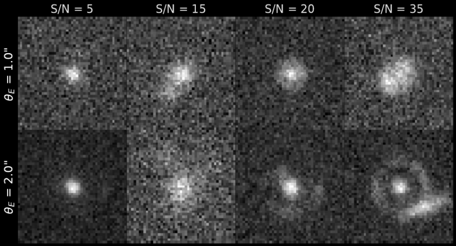
- Deep residual network (46 layers) with pre-activated bottleneck residual units

Lanusse et al. (2017)

CMU DeepLens: deep residual learning for strong lens finding

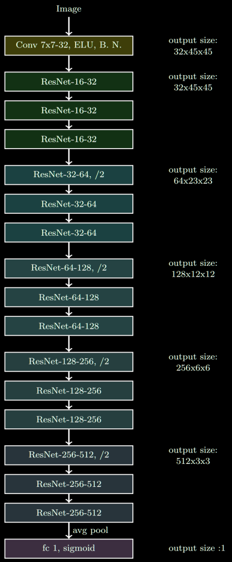


- Deep residual network (46 layers) with pre-activated bottleneck residual units
- Training on simulated LSST lenses:

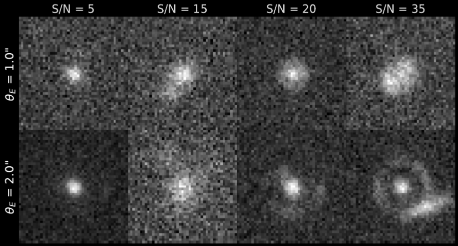


Lanusse et al. (2017)

CMU DeepLens: deep residual learning for strong lens finding



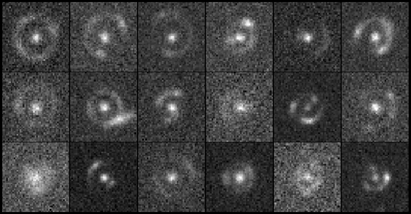
- Deep residual network (46 layers) with pre-activated bottleneck residual units
- Training on simulated LSST lenses:



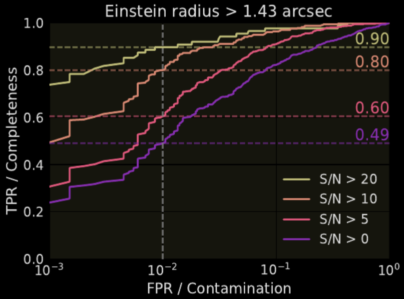
- Classification of 45x45 images in 350 μs $\Longrightarrow$ 9 hours to classify a sample of 10^8 lens candidates on a single GPU (Nvidia Titan X)

Lanusse et al. (2017)

performance on simulations



Highest probability lenses



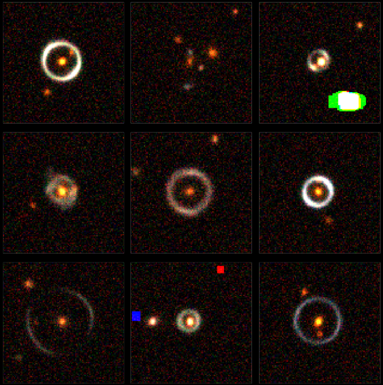
$$\text{True Positive Rate} = \frac{TP}{TP + FN}$$

$$\text{False Positive Rate} = \frac{FP}{FP + TN}$$

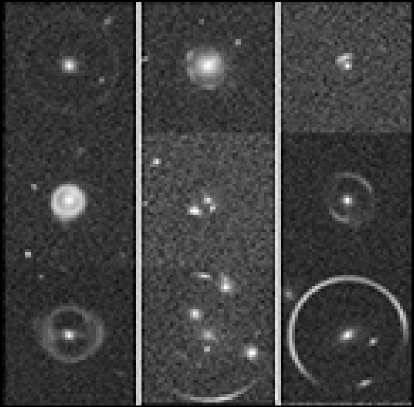
- TP: True Positives
- FN: False Negatives

- FP: False Positives
- TN: True Negatives

Euclid strong lens finding challenge

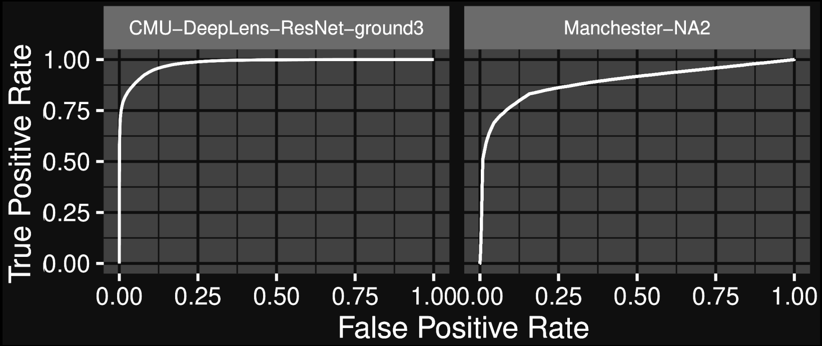


Ground based simulations



Space based simulations

Euclid strong lens finding challenge



- `CMU DeepLens` wins over 24 other methods (including other CNN methods) in space and ground challenge.
- Significantly outperforms human classification accuracy.

takeaway message

- An example of Deep Learning allowing us to handle the volume and data rate of future surveys

takeaway message

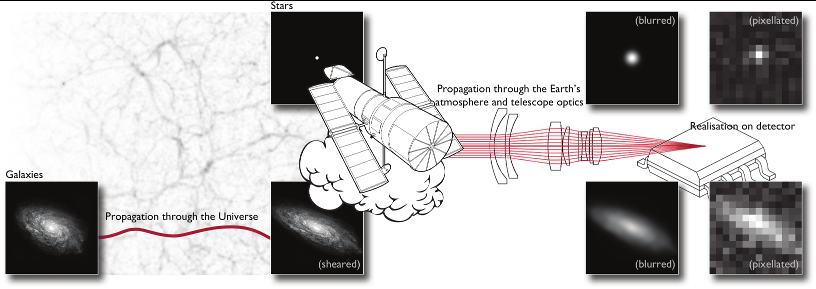
- An example of Deep Learning allowing us to handle the volume and data rate of future surveys
- Our automated lens finder is faster and more reliable than human volunteers.
    - Larger and more robust samples for the science analysis.

takeaway message

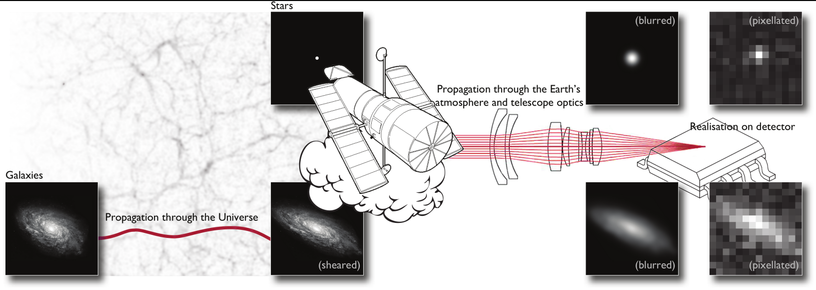
- An example of Deep Learning allowing us to handle the volume and data rate of future surveys
- Our automated lens finder is faster and more reliable than human volunteers.
    - Larger and more robust samples for the science analysis.
- Other uses for automated classification methods in LSST: photometric time series classification with recurrent neural networks

Deep Generative Models for modeling galaxy morphologies

the weak lensing shape measurement problem



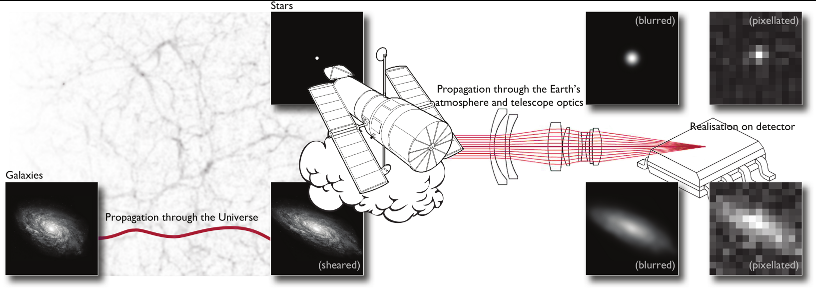
the weak lensing shape measurement problem



Shape measurement biases

$$< e > \quad = \quad (1 + m)\, \gamma + c$$

the weak lensing shape measurement problem

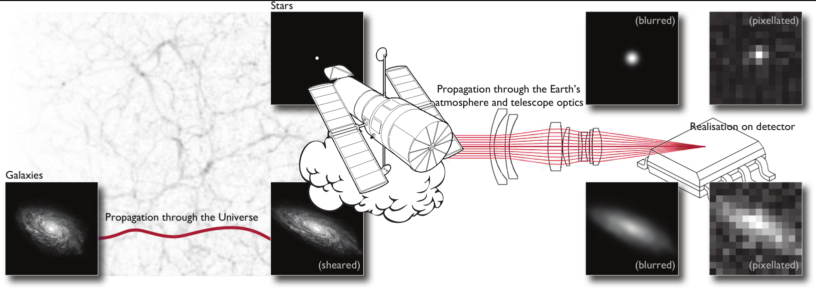


Shape measurement biases

$$< e > \quad = \quad (1 + m)\, \gamma + c$$

- Can be calibrated on image simulations

the weak lensing shape measurement problem



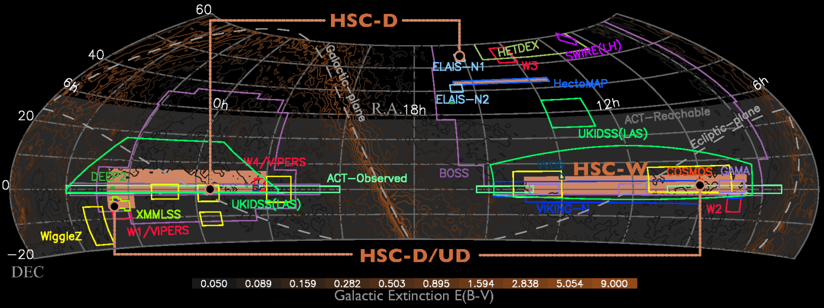
Shape measurement biases

$$< e > \quad = \quad (1+m)\,\gamma + c$$

- Can be calibrated on image simulations
- How complex do the simulations need to be ?

image simulations for shear calibration in the HSC survey

Mandelbaum, **Lanusse**, Leauthaud, Armstrong, et al. (2017)

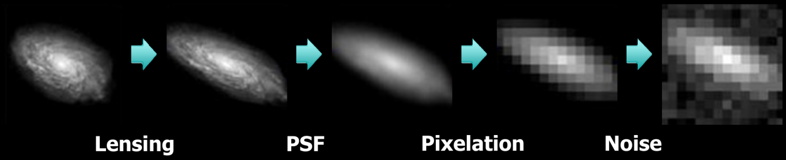


Hyper Suprime-Cam Subaru Strategic Program

- 1400 sq. deg. in the wide survey down to 26.4 mag in i-band
- Essentially a smaller area precursor of LSST

image simulations for shear calibration in the HSC survey

Mandelbaum, **Lanusse**, Leauthaud, Armstrong, et al. (2017)



The GREAT3 approach

- Input galaxies from deep HST/ACS COSMOS images (25.2 imag)
- Apply a range of PSFs and noise levels sampled from the survey
- Measure response of shape measurement to a known shear

image simulations for shear calibration in the HSC survey

Mandelbaum, **Lanusse**, Leauthaud, Armstrong, et al. (2017)

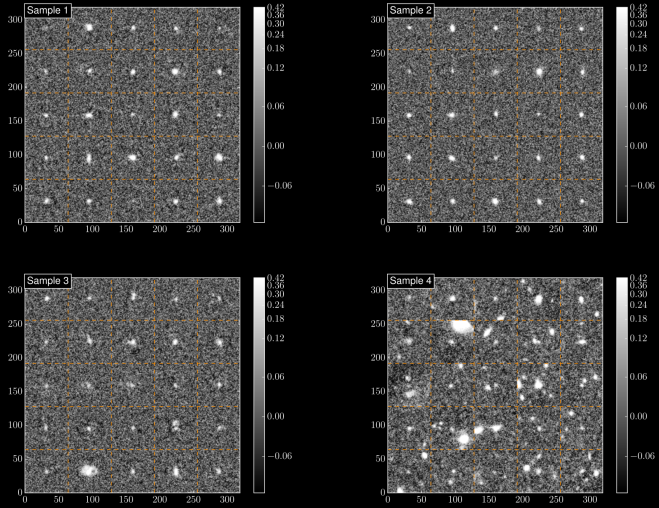
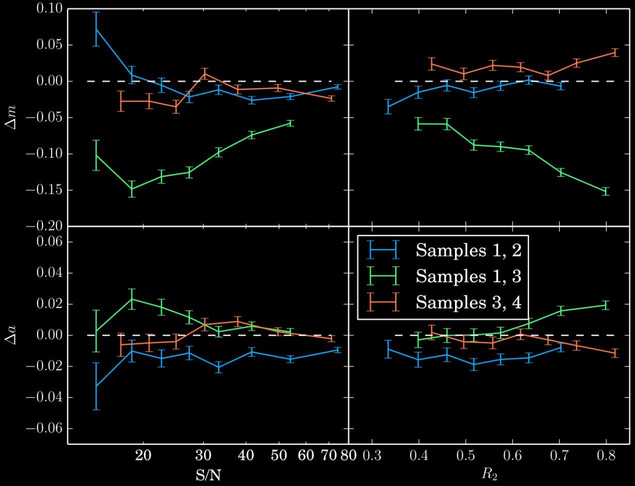
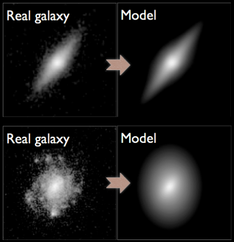


image simulations for shear calibration in the HSC survey

Mandelbaum, **Lanusse**, Leauthaud, Armstrong, et al. (2017)

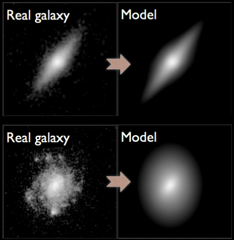


impact of galaxy morphology

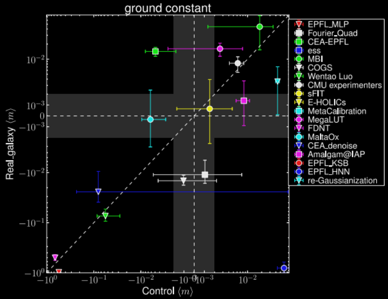


Mandelbaum et al. (2013)

impact of galaxy morphology

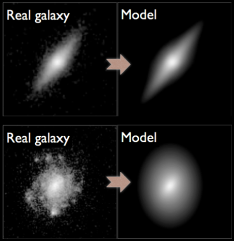


Mandelbaum et al. (2013)

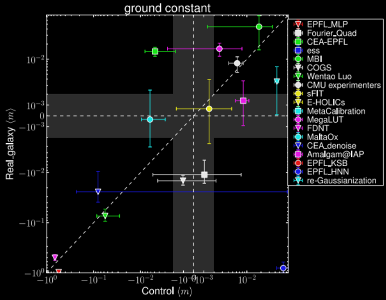


Mandelbaum et al. (2014)

impact of galaxy morphology



Mandelbaum et al. (2013)



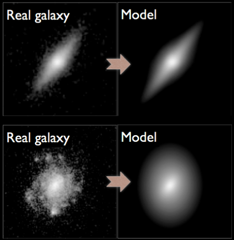
Mandelbaum et al. (2014)

The need for data-driven generative models

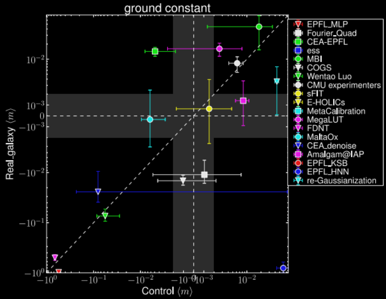
There can be two situations:

- Lack or inadequacy of physical model

impact of galaxy morphology



Mandelbaum et al. (2013)



Mandelbaum et al. (2014)

The need for data-driven generative models

There can be two situations:

- Lack or inadequacy of physical model
- Extremely computationally expensive simulations

Can we learn a model for the signal from the data itself ?

the evolution of deep generative models

- Deep Belief Network
  (Hinton et al. 2006)

the evolution of deep generative models

- Deep Belief Network
  (Hinton et al. 2006)

- Variational AutoEncoder
  (Kingma & Welling 2014)

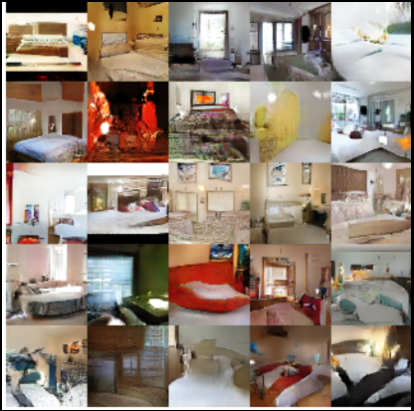
Credit: Alec Radford
`https://youtu.be/XNZIN7Jh3Sg`

the evolution of deep generative models

- Deep Belief Network
  (Hinton et al. 2006)

- Variational AutoEncoder
  (Kingma & Welling 2014)

- Generative Adversarial
  Network
  (Goodfellow et al. 2014)



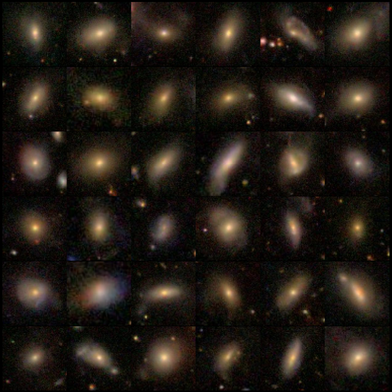
Deep Convolutional GAN (Radford et al. 2016)

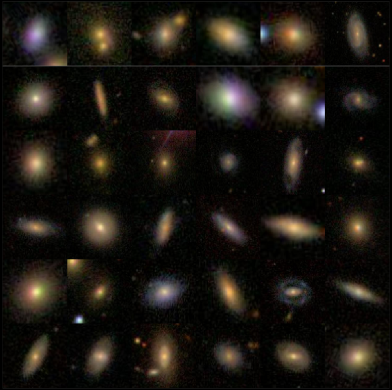
the evolution of deep generative models

- Deep Belief Network (Hinton et al. 2006)
- Variational AutoEncoder (Kingma & Welling 2014)
- Generative Adversarial Network (Goodfellow et al. 2014)
- Wasserstein GAN (Arjovsky et al. 2017)

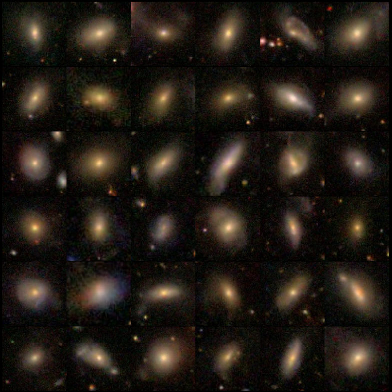
Progressive growing of wGAN-GP (Karras et al. 2017)

visual Turing test

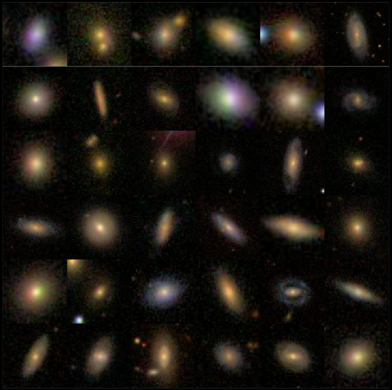




visual Turing test

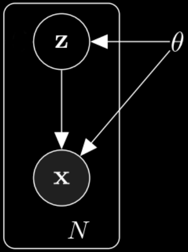


Mock - PixelCNN



Real - Sloan Digital Sky Survey

Auto-Encoding Variational Bayes (Kingma & Welling 2014)



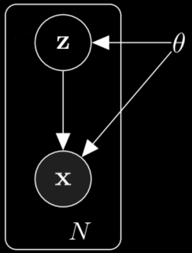
Generative model

Dataset of N i.i.d. samples $\{x_i\}$ generated from

$$x \sim p_{\theta_*}(x|z)\; p_{\theta_*}(z)$$

- z is a set of latent variables
- p_θ is some parametric distribution
- θ_* are the true model parameters

Auto-Encoding Variational Bayes (Kingma & Welling 2014)



Generative model

Dataset of N i.i.d. samples $\{x_i\}$ generated from

$$x \sim p_{\theta_*}(x|z)\; p_{\theta_*}(z)$$

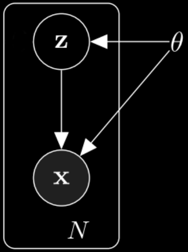
- z is a set of latent variables
- p_θ is some parametric distribution
- θ_* are the true model parameters

- In practice $p_\theta(x|z)$ can be parameterized with neural networks, e.g.:

$$p_\theta(x|z) = \mathcal{N}(x|\; \mu_\theta(z), \sigma^2_\theta(z))$$

where μ_θ and σ^2_θ are deep neural networks, with weights θ.

Auto-Encoding Variational Bayes (Kingma & Welling 2014)



Generative model

Dataset of N i.i.d. samples $\{x_i\}$ generated from

$$x \sim p_{\theta_*}(x|z)\ p_{\theta_*}(z)$$

- z is a set of latent variables
- p_θ is some parametric distribution
- θ_* are the true model parameters

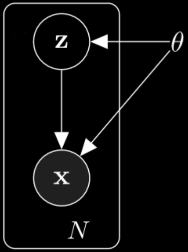
- In practice $p_\theta(x|z)$ can be parameterized with neural networks, e.g.:

$$p_\theta(x|z) = \mathcal{N}(x|\ \mu_\theta(z), \sigma^2_\theta(z))$$

  where μ_θ and σ^2_θ are deep neural networks, with weights θ.

- Training the model amounts to estimating the parameters θ maximizing the marginal likelihood $p_\theta(x) = \int p_\theta(z) p_\theta(x|z) dz$.

Auto-Encoding Variational Bayes (Kingma & Welling 2014)



Generative model

Dataset of N i.i.d. samples $\{x_i\}$ generated from

$$x \sim p_{\theta_*}(x|z)\, p_{\theta_*}(z)$$

- z is a set of latent variables
- p_θ is some parametric distribution
- θ_* are the true model parameters

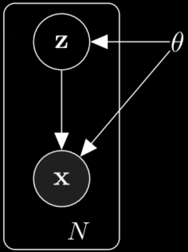
- In practice $p_\theta(x|z)$ can be parameterized with neural networks, e.g.:

  $$p_\theta(x|z) = \mathcal{N}(x|\ \mu_\theta(z), \sigma^2_\theta(z))$$

  where μ_θ and σ^2_θ are deep neural networks, with weights θ.

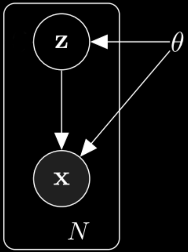
- Training the model amounts to estimating the parameters θ maximizing the marginal likelihood $p_\theta(x) = \int p_\theta(z) p_\theta(x|z) dz$.
  $\Longrightarrow$ Intractable and/or impractical in most situations

recognition model and variational lower bound



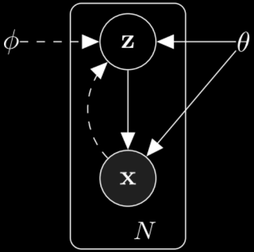
- Efficiently estimating $p_\theta(x) \approx \frac{1}{K}\sum_{k=1}^{K} p_\theta(x|z_k)$ requires sampling z_k where $p_\theta(x|z_k) \not\approx 0$.

recognition model and variational lower bound



- Efficiently estimating $p_\theta(x) \approx \frac{1}{K}\sum_{k=1}^{K} p_\theta(x|z_k)$ requires sampling z_k where $p_\theta(x|z_k) \not\approx 0$.
- The true posterior $p_\theta(z|x)$ is intractable.

recognition model and variational lower bound

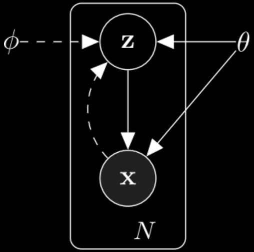


- Efficiently estimating $p_\theta(x) \approx \frac{1}{K}\sum_{k=1}^{K} p_\theta(x|z_k)$ requires sampling z_k where $p_\theta(x|z_k) \not\approx 0$.
- The true posterior $p_\theta(z|x)$ is intractable.

The auto-encoding solution

Introduce an auxiliary inference model $q_\phi(z|x)$ to approximate the true posterior.

recognition model and variational lower bound



- Efficiently estimating $p_\theta(x) \approx \frac{1}{K}\sum_{k=1}^{K} p_\theta(x|z_k)$ requires sampling z_k where $p_\theta(x|z_k) \not\approx 0$.
- The true posterior $p_\theta(z|x)$ is intractable.

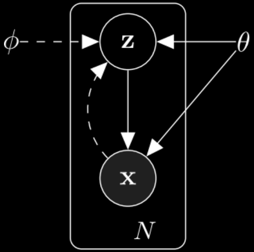
The auto-encoding solution

Introduce an auxiliary inference model $q_\phi(z|x)$ to approximate the true posterior.

- Working out the Kullback-Leibler divergence between $q_\phi(z|x)$ and $p_\theta(z|x)$ yields:

$$\log p_\theta(x) \; - \; \mathbb{D}_{\text{KL}}(q_\phi(z|x) \| p_\theta(z|x)) \; = \; \mathbb{E}_{z \sim q_\phi(\cdot|x)}[\log p_\theta(x|z)] \; - \; \mathbb{D}_{\text{KL}}(q_\phi(z|x) \| p_\theta(z))$$

recognition model and variational lower bound



- Efficiently estimating $p_\theta(x) \approx \frac{1}{K}\sum_{k=1}^{K} p_\theta(x|z_k)$ requires sampling z_k where $p_\theta(x|z_k) \not\approx 0$.
- The true posterior $p_\theta(z|x)$ is intractable.

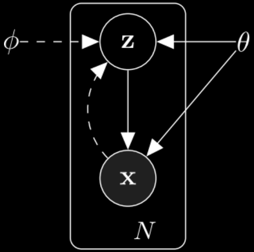
The auto-encoding solution

Introduce an auxiliary inference model $q_\phi(z|x)$ to approximate the true posterior.

- Working out the Kullback-Leibler divergence between $q_\phi(z|x)$ and $p_\theta(z|x)$ yields:

$$\log p_\theta(x) - \underbrace{\mathbb{D}_{\mathrm{KL}}(q_\phi(z|x) \| p_\theta(z|x))}_{\text{Unknown, but } > 0} = \mathbb{E}_{z\sim q_\phi(\cdot|x)}[\log p_\theta(x|z)] - \mathbb{D}_{\mathrm{KL}}(q_\phi(z|x) \| p_\theta(z))$$

recognition model and variational lower bound



- Efficiently estimating $p_\theta(x) \approx \frac{1}{K}\sum_{k=1}^{K} p_\theta(x|z_k)$ requires sampling z_k where $p_\theta(x|z_k) \not\approx 0$.
- The true posterior $p_\theta(z|x)$ is intractable.

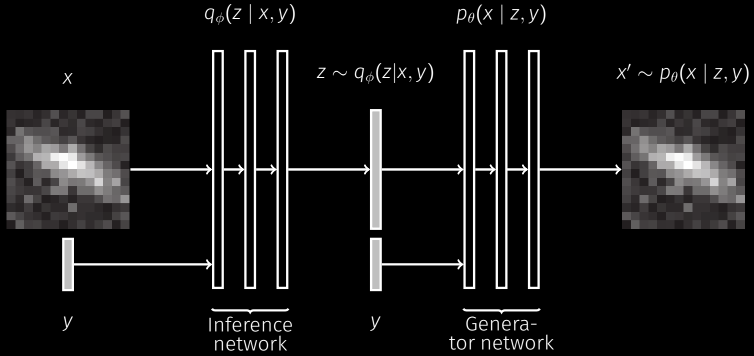
The auto-encoding solution

Introduce an auxiliary inference model $q_\phi(z|x)$ to approximate the true posterior.

- Working out the Kullback-Leibler divergence between $q_\phi(z|x)$ and $p_\theta(z|x)$ yields:

$$\log p_\theta(x) - \underbrace{\mathbb{D}_{\mathrm{KL}}(q_\phi(z|x)\|p_\theta(z|x))}_{\text{Unknown, but } > 0} = \underbrace{\mathbb{E}_{z\sim q_\phi(\cdot|x)}[\log p_\theta(x|z)] - \mathbb{D}_{\mathrm{KL}}(q_\phi(z|x)\|p_\theta(z))}_{\text{Differentiable lower bound on the likelihood}}$$

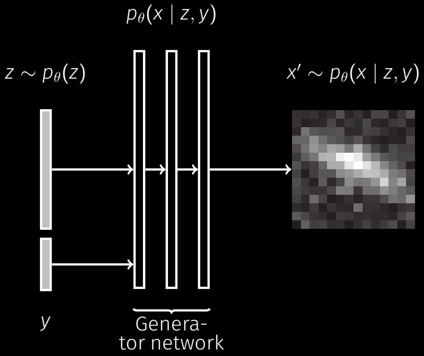
Conditional Variational AutoEncoder (CVAE)



Ravanbakhsh, **Lanusse** et al. (2017)

$$\log(p_\theta(x \mid y)) \geq -\underbrace{\mathbb{D}_{\mathrm{KL}}(q_\phi(z \mid x, y) \| p_\theta(z))}_{\text{Code regularisation}} + \underbrace{\mathbb{E}_{z \sim q_\phi(\cdot|x,y)}[\log p_\theta(x \mid z, y)]}_{\text{Reconstruction error}}$$

Conditional Variational AutoEncoder (CVAE)



Ravanbakhsh, **Lanusse** et al. (2017)

modeling galaxy images from the Hubble Space Telescope

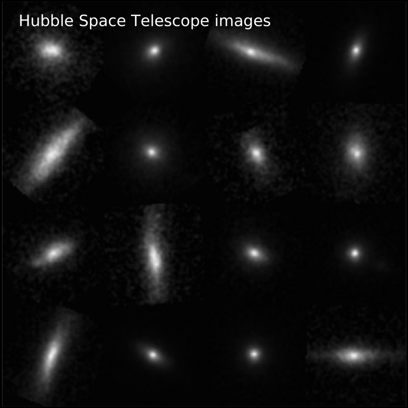
Experimental setup

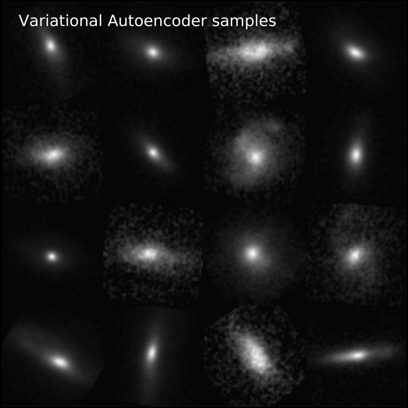
- Training set: postage stamps from COSMOS HST/ACS survey
- Conditional model: Half-light radius, brightness, redshift

modeling galaxy images from the Hubble Space Telescope

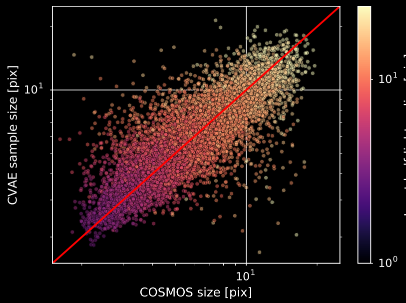
Experimental setup

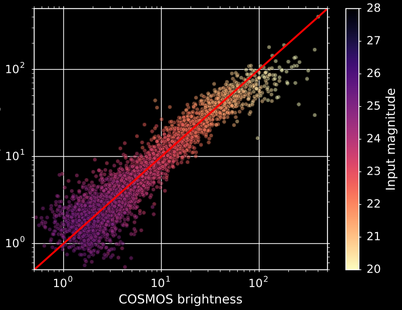
- Training set: postage stamps from COSMOS HST/ACS survey
- Conditional model: Half-light radius, brightness, redshift



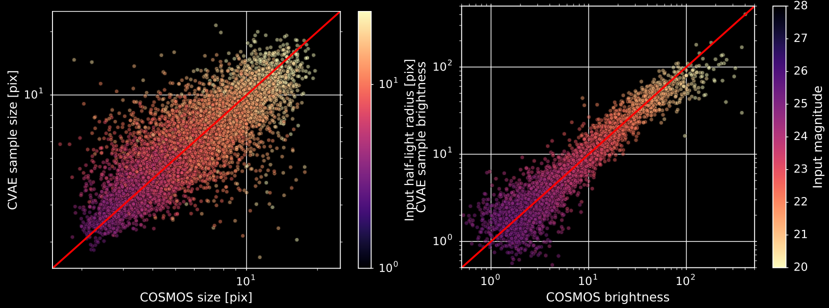


testing the conditional generation



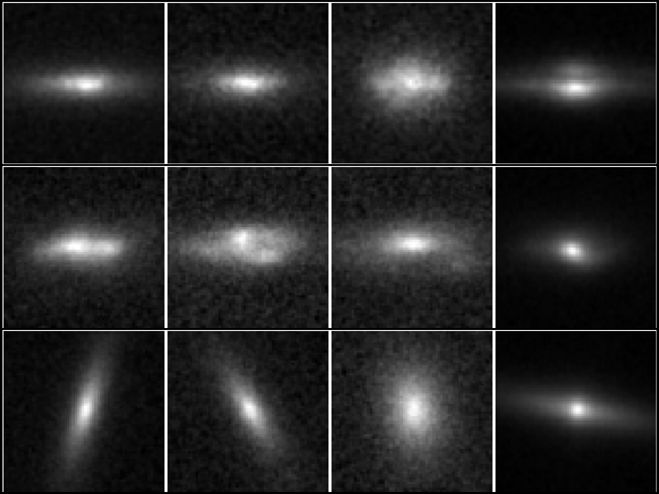


testing the conditional generation



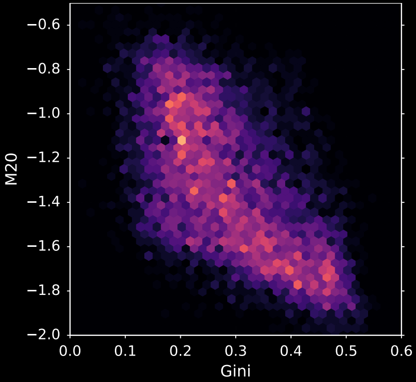
$\Longrightarrow$ We can generate galaxies of desired size and brightness

morphological statistics

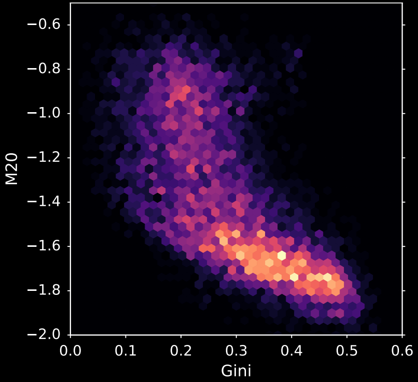


From top to bottom: Real COSMOS galaxies, CVAE samples, Parametric fits

morphological statistics

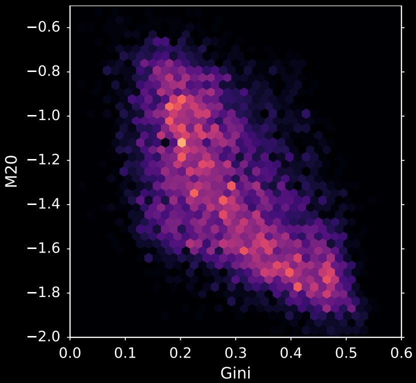


COSMOS images

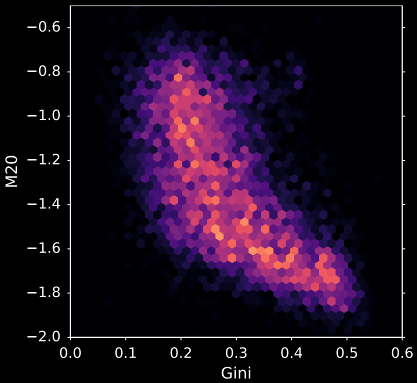


Parametric model

morphological statistics



COSMOS images



CVAE samples

takeaway message

- Generative models are a data driven way of completing our physical modeling

takeaway message

- Generative models are a data driven way of completing our physical modeling
- Implementing these models inside the `GalSim` simulation software used to simulate the LSST survey
    - Adversarial code manipulation
    - Properly handling COSMOS PSF and correlated noise

takeaway message

- Generative models are a data driven way of completing our physical modeling
- Implementing these models inside the `GalSim` simulation software used to simulate the LSST survey
  - Adversarial code manipulation
  - Properly handling COSMOS PSF and correlated noise
- Will allow for an extra degree of realism in LSST simulations
  $\Longrightarrow$ Essential to the calibration of science pipeline

takeaway message

- Generative models are a data driven way of completing our physical modeling

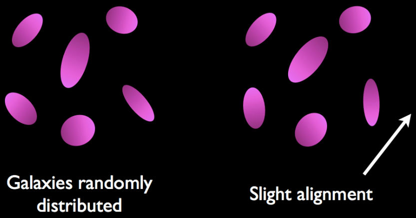
- Implementing these models inside the `GalSim` simulation software used to simulate the LSST survey
  - Adversarial code manipulation
  - Properly handling COSMOS PSF and correlated noise

- Will allow for an extra degree of realism in LSST simulations
  $\Longrightarrow$ Essential to the calibration of science pipeline

- Other uses for generative models: any time a complex image prior is required (denoising, deconvolution, deblending, etc)

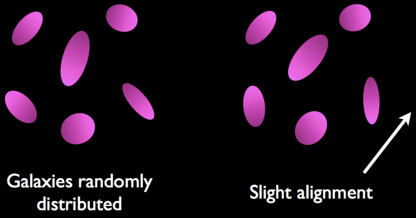
Graph convolutional networks for cosmological simulations

intrinsic alignment of galaxies



$\epsilon = \epsilon_i + \gamma$ with $< \epsilon_i >= 0$

intrinsic alignment of galaxies



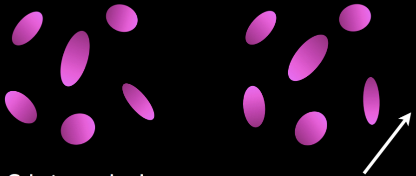
$\epsilon = \epsilon_i + \gamma$ with $\underbrace{< \epsilon_i >= 0}_{\text{not completely true}}$

intrinsic alignment of galaxies

$\epsilon = \epsilon_i + \gamma$ with $\underbrace{< \epsilon_i > = 0}_{\text{not completely true}}$

$$\underbrace{< \epsilon\epsilon' >}_{\text{measured}} = \underbrace{< \gamma\gamma' >}_{\text{cosmological signal}} + \underbrace{< \epsilon_i\epsilon_i' >}_{\text{II}} + \underbrace{< \gamma\epsilon_i' > + < \epsilon_i\gamma' >}_{\text{GI}}$$

intrinsic alignment of galaxies

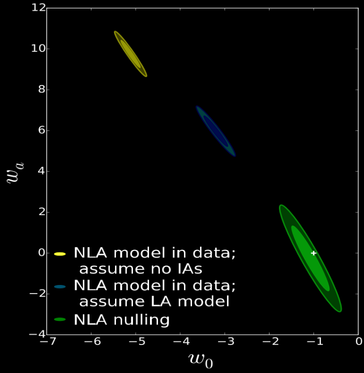


Galaxies randomly distributed

Slight alignment

$\epsilon = \epsilon_i + \gamma$ with $\underbrace{< \epsilon_i >= 0}_{\text{not completely true}}$

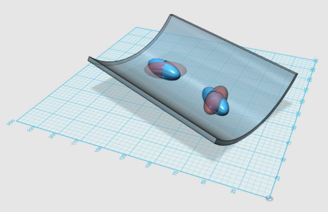
Impact on dark energy constraints



Kirk et al. (2015)

$$\underbrace{< \epsilon\epsilon' >}_{\text{measured}} = \underbrace{< \gamma\gamma' >}_{\text{cosmological signal}} + \underbrace{< \epsilon_i\epsilon_i' >}_{\text{II}} + \underbrace{< \gamma\epsilon_i' > + < \epsilon_i\gamma' >}_{\text{GI}}$$

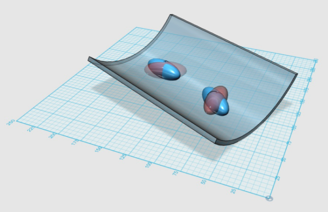
why does this happen ?



Kiessling et al. (2015)

- Tidal interactions with local gravitational potential
  $\Longrightarrow$ Can be analytically modeled on large scales

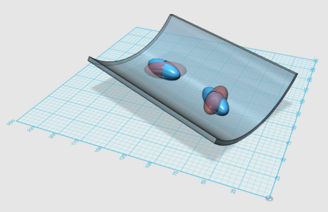
why does this happen ?



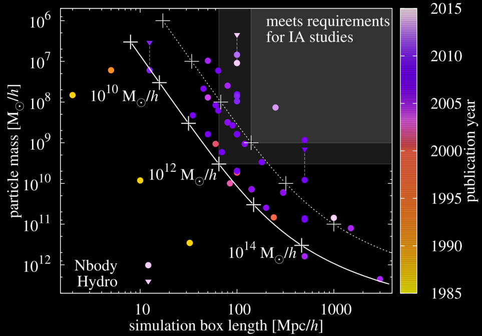
Kiessling et al. (2015)

- Tidal interactions with local gravitational potential
  $\Longrightarrow$ Can be analytically modeled on large scales
- Much more complicated in details, no single model for all galaxy types, impacted by baryonic physics

why does this happen ?



Kiessling et al. (2015)

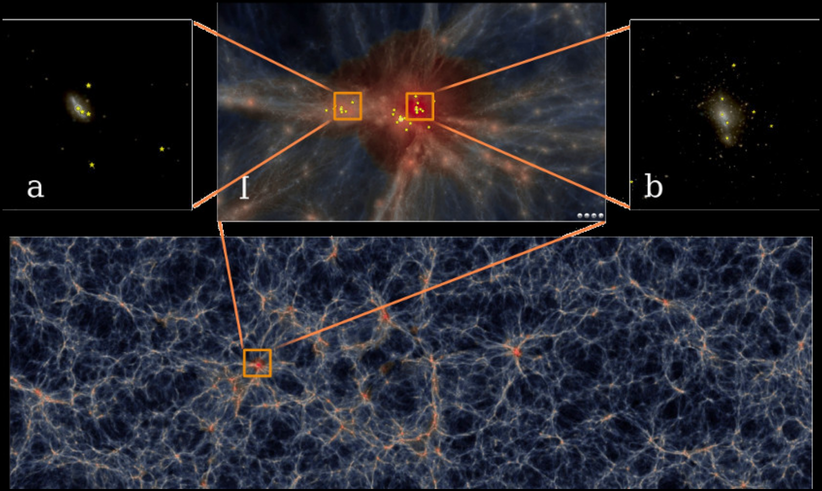


- Tidal interactions with local gravitational potential
  $\Longrightarrow$ Can be analytically modeled on large scales

- Much more complicated in details, no single model for all galaxy types, impacted by baryonic physics
  $\Longrightarrow$ Study requires expensive hydrodynamical simulation

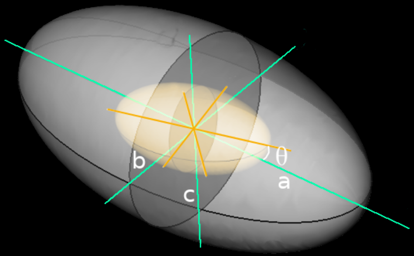
How to produce mock galaxy catalogs on large cosmological volumes with realistic alignments ?

Massive Black II simulation (Khandai et al, 2015)

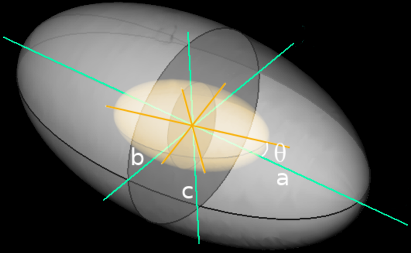


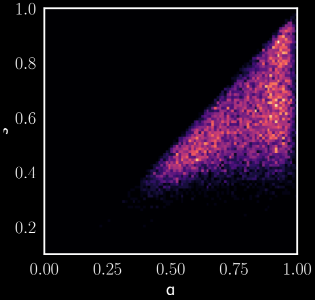
- $100h^{-1}$Mpc hydrodynamical simulation ran to $z = 0.06$
- Corresponding dark matter only simulation

what quantities are we looking at ?



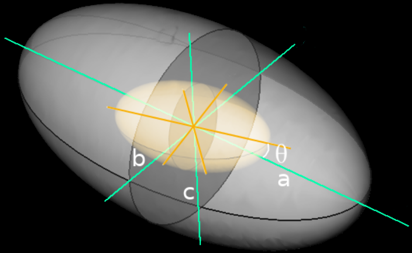
what quantities are we looking at ?

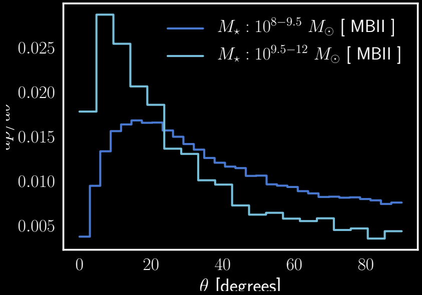




- Shape parameters: $q = \frac{b}{a}$ $\quad s = \frac{c}{a}$

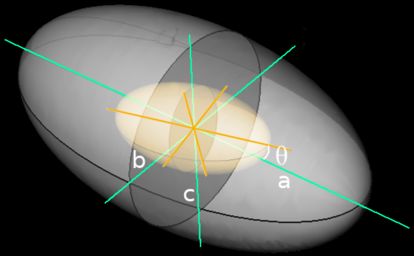
what quantities are we looking at ?

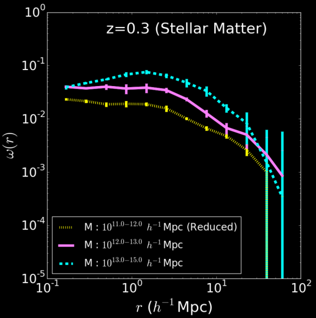




- Shape parameters: $q = \frac{b}{a}$ $\qquad s = \frac{c}{a}$
- Misalignment angle: $\theta = \text{arccos}\left(|\hat{e}_a^{(dm)} \cdot \hat{e}_g^{(gal)}|\right)$

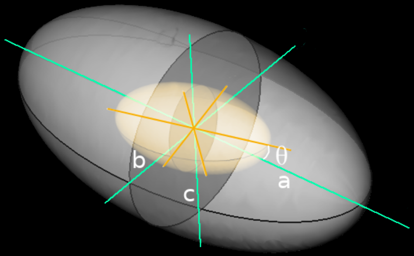
what quantities are we looking at ?

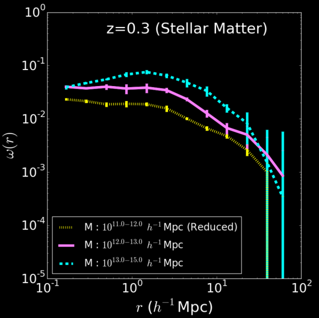




- Shape parameters: $q = \frac{b}{a}$ $\quad s = \frac{c}{a}$
- Misalignment angle: $\theta = \arccos\left(|\hat{e}_a^{(dm)} \cdot \hat{e}_g^{(gal)}|\right)$
- Ellipticity-Direction correlation function: $\omega(r) =< |\hat{e}_a \cdot \hat{\mathbf{r}}|^2 > -\frac{1}{3}$

what quantities are we looking at ?

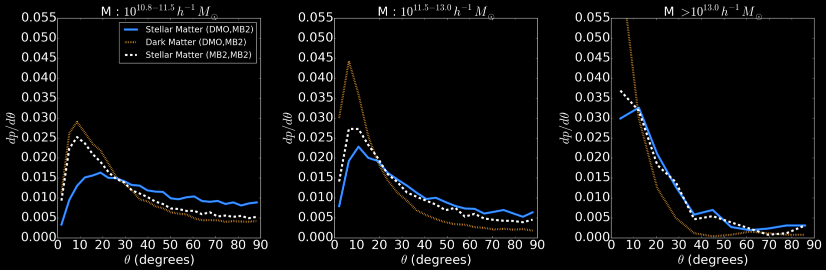




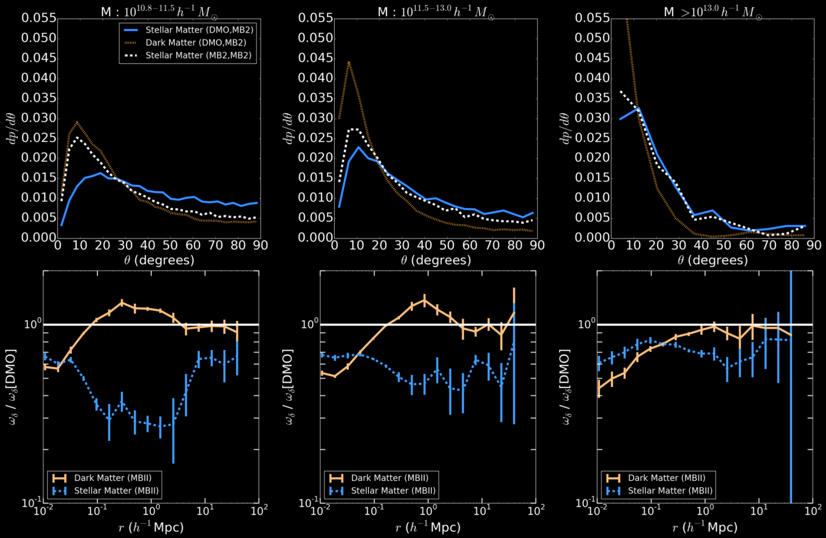
- Shape parameters: $q = \frac{b}{a}$ $\quad s = \frac{c}{a}$
- Misalignment angle: $\theta = \arccos\left(|\hat{e}_a^{(dm)} \cdot \hat{e}_g^{(gal)}|\right)$
- Ellipticity-Direction correlation function: $\omega(r) =< |\hat{e}_a \cdot \hat{\mathbf{r}}|^2 > -\frac{1}{3}$

$\Longrightarrow$ Depends on redshift, galaxy types, stellar mass, baryons, merger history...

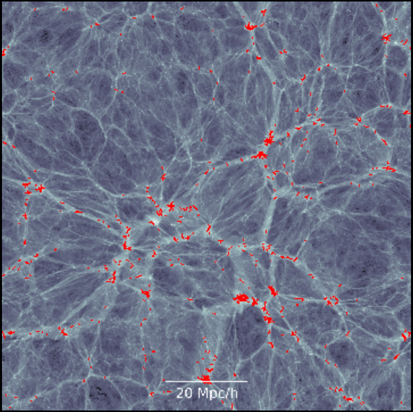
example: impact of baryons (Tenneti et al. 2015)



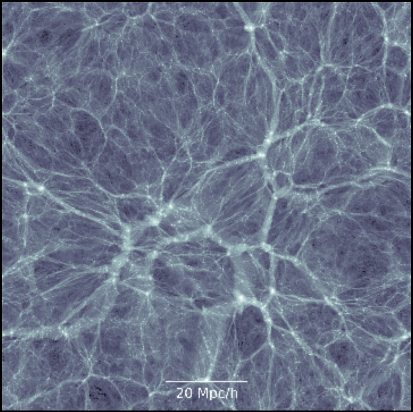
example: impact of baryons (Tenneti et al. 2015)



inpainting intrinsic aligments on N-body simulations



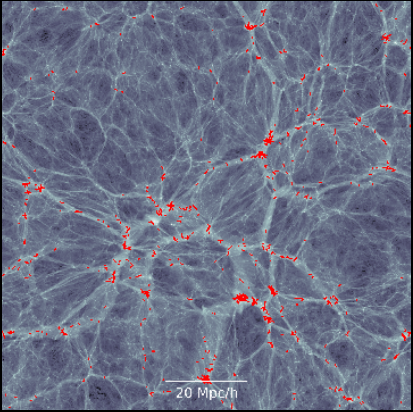
Massive Black II



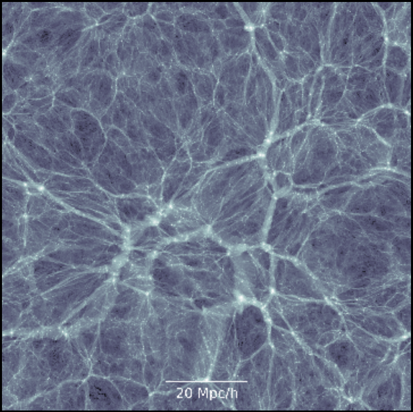
Dark Matter Only

Image credit: Tenneti et al. (2015)

inpainting intrinsic aligments on N-body simulations



Massive Black II

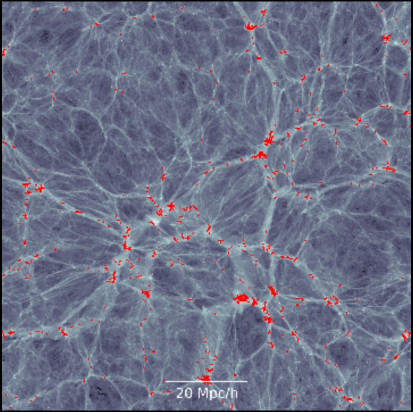


Dark Matter Only

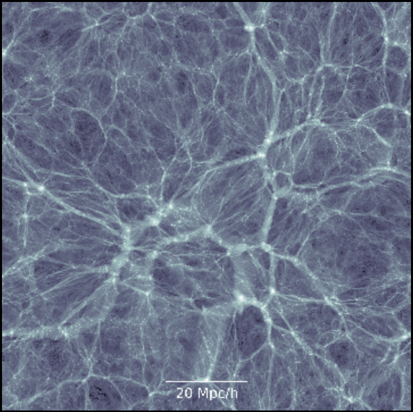
Image credit: Tenneti et al. (2015)

$$\mathrm{gal} \sim p\,(\textit{stellar properties} \mid \textit{dark matter properties})$$

inpainting intrinsic aligments on N-body simulations



Massive Black II

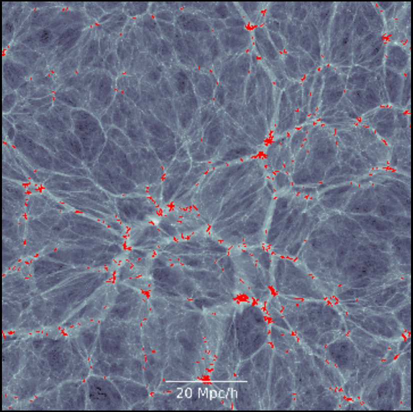


Dark Matter Only

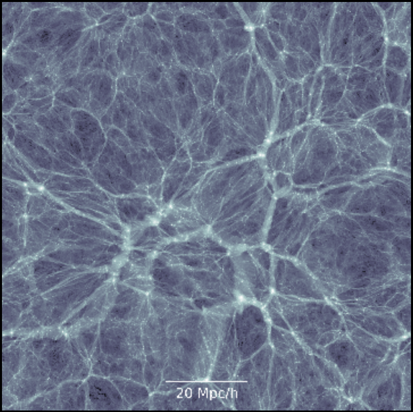
Image credit: Tenneti et al. (2015)

$$\mathrm{gal}_i \quad \sim \quad p\,(M_*,\ \theta,\ q_\star,\ s_\star,\ \ldots |\, M_{DM,i}, V_{disp,i}, q_{DM,i}, s_{DM,i}, \ldots)$$

inpainting intrinsic aligments on N-body simulations



Massive Black II



Dark Matter Only

Image credit: Tenneti et al. (2015)

$$\mathrm{gal}_i \quad \sim \quad p\left(M_*,\ \theta,\ q_\star,\ s_\star,\ \ldots \,|\, M_{DM,i}, V_{disp,i}, q_{DM,i}, s_{DM,i}, \ldots\right)$$

$\Longrightarrow$ How to model this conditional distribution ?

first approach using a CVAE

- Use a CVAE to independently draw each galaxy

first approach using a CVAE

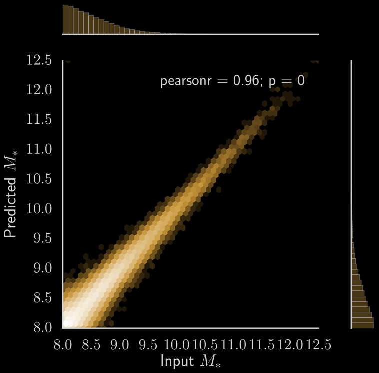
- Use a CVAE to independently draw each galaxy
- Conditioned on Dark Matter subhalo properties: mass, shape, orientation, local tidal field

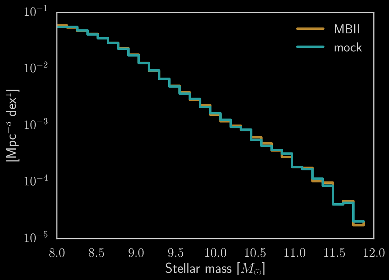
first approach using a CVAE

- Use a CVAE to independently draw each galaxy
- Conditioned on Dark Matter subhalo properties:
  mass, shape, orientation, local tidal field
- Predicts distribution of Stellar properties for each galaxy: stellar mass, shape, misalignment angles, bulge to total ratio

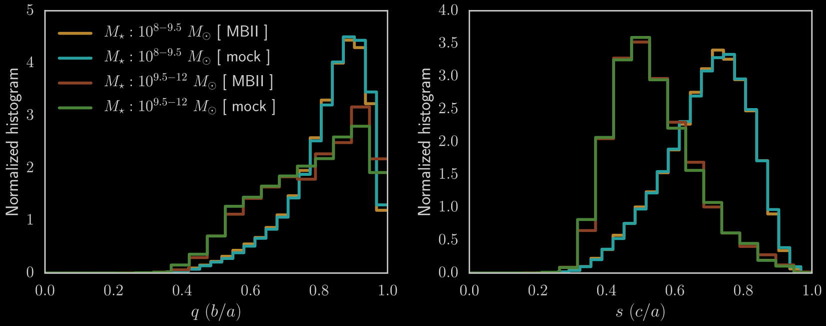
first approach using a CVAE

- Use a CVAE to independently draw each galaxy
- Conditioned on Dark Matter subhalo properties: mass, shape, orientation, local tidal field
- Predicts distribution of Stellar properties for each galaxy: stellar mass, shape, misalignment angles, bulge to total ratio

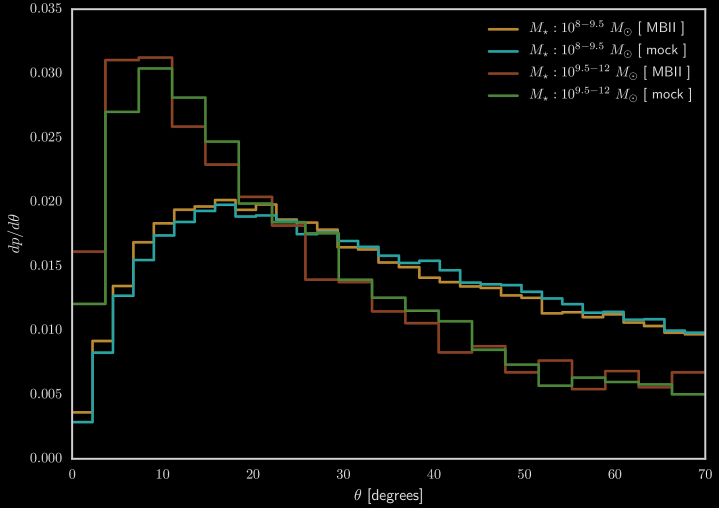




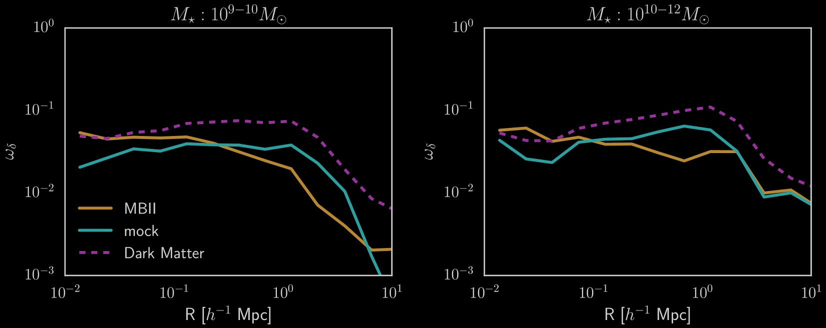
predicted galaxy shapes



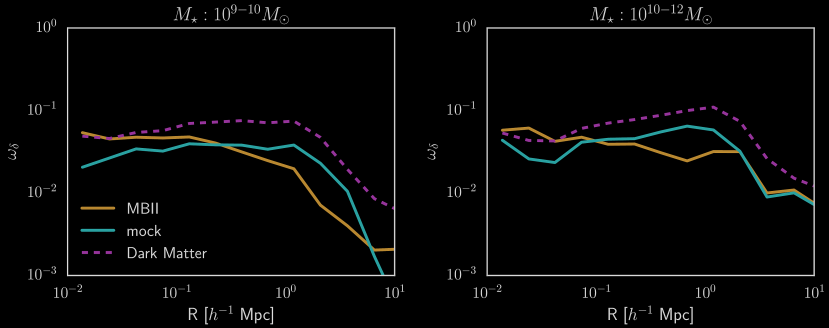
predicted misalignment angles



Ellipticity-Direction correlation function



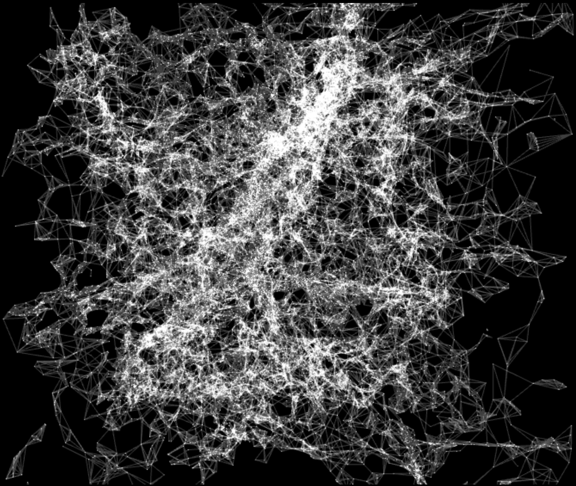
Ellipticity-Direction correlation function



Fails to reproduce 2pt functions

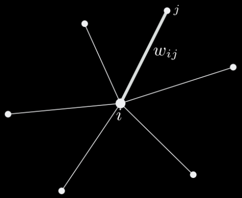
- The galaxies are randomly misaligned
  $\rightarrow$ suppresses power evenly on all scales
- To reproduce 2pt functions, galaxies must point in a specific direction with respect to their neighbors

the cosmic web as a graph



MB II simulation, animation credit: Kim Albrecht

spectral theory on graphs



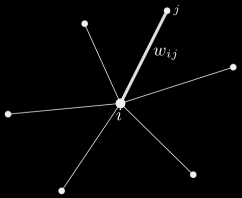
Graph Fourier Transform

Defined by the eigendecomposition of the graph Laplacian $\mathbf{L} = \mathbf{U}\mathbf{U}^T$

The graph Fourier transform of a signal f is

$$\hat{f} = \mathbf{U}^T f$$

spectral theory on graphs



Graph Fourier Transform

Defined by the eigendecomposition of the graph Laplacian $\mathbf{L} = \mathbf{U}\mathbf{U}^T$

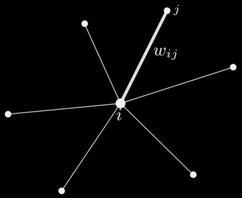
The graph Fourier transform of a signal f is

$$\hat{f} = \mathbf{U}^T f$$

- The convolution of a signal f with a filter g_θ can be defined as

$$\boxed{g_\theta \star f = \mathbf{U}\ \hat{g}_\theta\ \mathbf{U}^T f}$$

spectral theory on graphs



Graph Fourier Transform

Defined by the eigendecomposition of the graph Laplacian $\mathbf{L} = \mathbf{U}\mathbf{U}^T$

The graph Fourier transform of a signal f is

$$\hat{f} = \mathbf{U}^T f$$

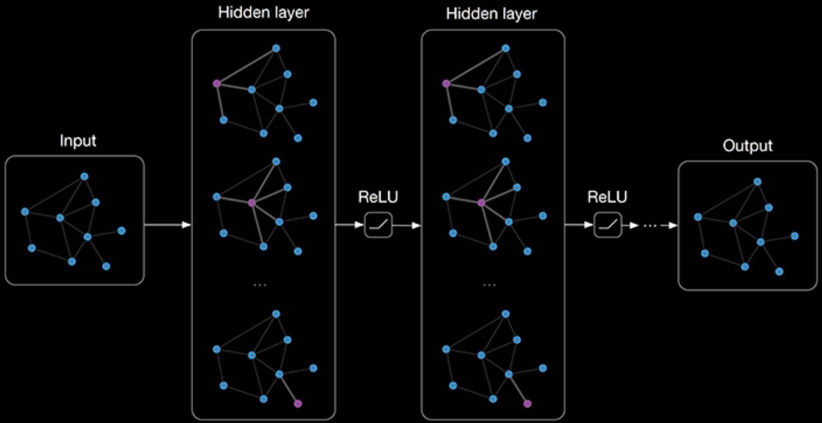
- The convolution of a signal f with a filter g_θ can be defined as

$$\boxed{g_\theta \star f = \mathbf{U}\, \hat{g}_\theta\, \mathbf{U}^T f}$$

Spectral graph convolutions are expensive

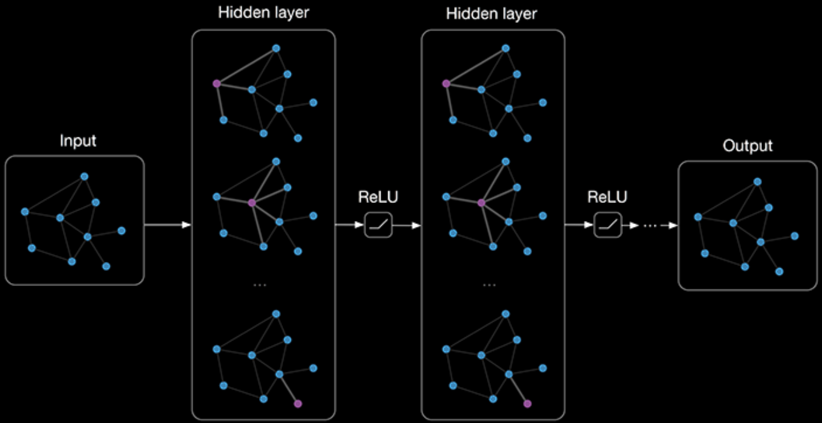
- Computing the graph Fourier transform of $\mathbf{U}$ is expensive for large graphs
- Computing one convolution requires $\mathcal{O}(N^2)$ operations

deep learning on graphs (Kipf & Welling (2017))



- Convolution layer reduces to: $g_{\theta'} \star f = \theta'_0 f - \theta'_1 \, D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \, f$

deep learning on graphs (Kipf & Welling (2017))



- Convolution layer reduces to: $g_{\theta'} \star f = \theta'_0 f - \theta'_1 \, D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \, f$
- Tractable graph convolutions using an approximation restricted to first neighbors

preliminary results with new model

- Model adapts mixture density networks and deep residual networks to graphs

preliminary results with new model

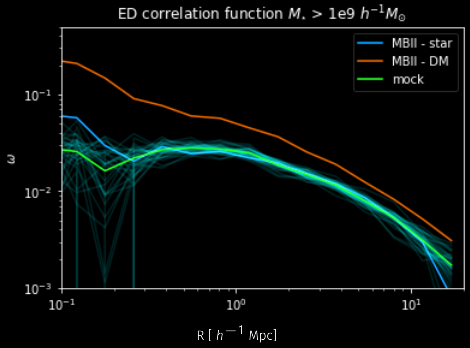
- Model adapts mixture density networks and deep residual networks to graphs
- Conditioned on Dark Matter subhalo properties of all galaxies in the graph: mass, shape, local tidal field

preliminary results with new model

- Model adapts mixture density networks and deep residual networks to graphs
- Conditioned on Dark Matter subhalo properties of all galaxies in the graph: mass, shape, local tidal field
- Predicts 3D orientations of all galaxies in the graph.

preliminary results with new model

- Model adapts mixture density networks and deep residual networks to graphs
- Conditioned on Dark Matter subhalo properties of all galaxies in the graph: mass, shape, local tidal field
- Predicts 3D orientations of all galaxies in the graph.



takeaway message

- Exciting new framework to empirically populate large volume simulations with realistic galaxy populations
  - Easily extendable to include merger tree information
  - Extend/complement existing empirical/semi-analytical model

takeaway message

- Exciting new framework to empirically populate large volume simulations with realistic galaxy populations
  - Easily extendable to include merger tree information
  - Extend/complement existing empirical/semi-analytical model

- Will add to the realism of cosmological simulations and allow us to test IA mitigation techniques.
  - Being implemented as part of the simulation pipeline for the LSST DESC Second Data Challenge

takeaway message

- Exciting new framework to empirically populate large volume simulations with realistic galaxy populations
  - Easily extendable to include merger tree information
  - Extend/complement existing empirical/semi-analytical model
- Will add to the realism of cosmological simulations and allow us to test IA mitigation techniques.
  - Being implemented as part of the simulation pipeline for the LSST DESC Second Data Challenge
- Neural networks on graphs are powerful tools for working with non euclidean data.
  $\Longrightarrow$ See Bronstein et al. 2017 (arXiv:1611.08097) for a recent overview.

Conclusion

conclusion

What can deep learning do for cosmology ?

- Model and analyze large volumes of complex datasets

conclusion

What can deep learning do for cosmology ?

- Model and analyze large volumes of complex datasets
- Open new and powerful ways to look at the data
  - Image detection for finding rare astrophysical objects

conclusion

What can deep learning do for cosmology ?

- Model and analyze large volumes of complex datasets
- Open new and powerful ways to look at the data
  - Image detection for finding rare astrophysical objects
- Data driven way of complementing our physical models
  - Modeling realistic galaxy morphologies
  - Modeling galaxy properties in numerical simulations

$\Longrightarrow$ Goes towards improving the accuracy of our cosmology constraints.

Any questions?